

# Base station cooperation on the downlink: Large system analysis

Randa Zakhour and Stephen V. Hanly

**Abstract**—This paper considers maximizing the network-wide minimum supported rate in the downlink of a two-cell system, where each base station (BS) is endowed with multiple antennas. This is done for different levels of cell cooperation. At one extreme, we consider single cell processing where the BS is oblivious to the interference it is creating at the other cell. At the other extreme, we consider full cooperative macroscopic beamforming. In between, we consider coordinated beamforming, which takes account of inter-cell interference, but does not require full cooperation between the BSs. We combine elements of Lagrangian duality and large system analysis to obtain limiting SINRs and bit-rates, allowing comparison between the considered schemes. The main contributions of the paper are theorems which provide concise formulas for optimal transmit power, beamforming vectors, and achieved signal to interference and noise ratio (SINR) for the considered schemes. The formulas obtained are valid for the limit in which the number of users per cell,  $K$ , and the number of antennas per base station,  $N$ , tend to infinity, with fixed ratio  $\beta = K/N$ . These theorems also provide expressions for the effective bandwidths occupied by users, and the effective interference caused in the adjacent cell, which allow direct comparisons between the considered schemes.

**Index Terms**—linear precoding, regularized zero forcing, inter-cell interference, base station cooperation, multicell processing, cellular systems, MIMO broadcast channel, interference channel

## I. INTRODUCTION

### A. Problem scope

Consider the downlink (DL) of a cellular system in which a base station (BS) services many mobiles within the cell. If the BS is equipped with multiple antennas we have the classic MIMO broadcast channel (BC), which has been the focus of much attention in the past few years, including the recent, celebrated characterization of the capacity region using dirty paper coding [1]. There has also been substantial interest in devising suboptimal, but practical approaches based on linear precoding techniques (i.e. beamforming).

The MIMO BC is the appropriate model for a single isolated cell. What happens when we bring several cells together, so that each is affected by the interference from the others? This paper analyzes the performance of linear precoding in a multicell setting.

Consideration of interference leads us to examine the system level architectural issue of cooperation between nodes in the

network. The traditional approach to interference is to partition the cells in time or bandwidth to avoid a strong interference coupling between adjacent cells. However, with multiple antennas at each base station, this may be suboptimal. When multiple antennas are incorporated at each BS, we can tradeoff the maximization of the rates of the in-cell users (ignoring interference), with the minimization of the interference spilled over into the other cells. If enough cooperation is enabled, these two objectives can go hand in hand.

A large body of research has recently dealt with cooperation and coordination in multicell systems. Many papers are concerned with developing new algorithms, to meet various proposed performance objectives (e.g. transmit power minimization under given SINR constraints for coordinated beamforming (CBf) [2], [3], or minimum SINR maximization for network MIMO [4]). Others provide a performance analysis of a given scheme under a particular channel model.

One can distinguish between scenarios where BSs each serve a different group of users, and cases where all the transmitters jointly transmit to all users in the system, the so-called multicell processing (MCP), macrodiversity or network MIMO. For MCP, a classical model for performance analysis is Wyner's model [5], [6]. This model was first used on the DL in [7], where a linear pre-processing dirty paper coding approach is proposed. In [6], the sum rate is characterized for the case where single-antenna base stations pool their antennas together to perform zero-forcing (ZF) to the users in the system. More precisely, they consider a circular variant of the infinite linear Wyner model for both non-fading and fading scenarios, with scheduling based on local channel statistics. Results are derived for the regime in which the number of cooperating BSs tends to infinity, and scaling results are also obtained by letting the number of users per cell do the same.

In this work, we focus on optimizing linear precoding under different states of CSI and data sharing. We assume that each BS can simultaneously serve more than one user, and we formulate the problem of maximizing the minimum network-wide achievable rate, i.e. rate balancing, under the following three different architectures:

- i) Single cell processing (SCP), in which each BS has perfect CSI about mobiles in its cell, but no knowledge about the channels to mobiles in other cells. Here, the BS can control interference between the mobiles in its own cell, but not the interference spill-over to other cells.
- ii) CBf, in which the beamforming decisions at each BS take into account the impact of interference on the other cells. In this approach, each BS has CSI about the channels to its own users, but it also knows the channels

R. Zakhour is with the Electrical and Electronic Engineering Department, University of Melbourne, Australia: rzakhour@unimelb.edu.au, S. V. Hanly is with the Department of Electrical and Computer Engineering, National University of Singapore: elehsv@nus.edu.sg. This work was supported by the Australian Research Council (ARC) under grant DP0984862, and NUS grant R-263-000-572-133.

to the mobiles in the other cells. This allows the BS to control the interference it causes to the mobiles in other cells.

- iii) MCP, in which the BSs cooperate to jointly precode signals to all the mobiles in all cells. Through cooperation, the BSs know the channels from all BSs to all the mobiles in the whole network, and hence they can jointly beamform, using their combined resources.

In the first two approaches, user data is routed to a single BS only, whereas in the MCP case, it is routed to all cooperating BSs. We specialize our derivations to only two cells. Nevertheless, such an approach can be extended to larger networks, if one wishes to perform system-wide optimization of a large cellular network. This is the object of ongoing research [8].

Even the two cell model gets very complicated when we have multiple antennas at the BSs and many mobiles in each cell. With independent flat fading between each transmitter-receiver pair, many parameters need to be specified. For this reason, we assume that they are selected randomly, and we take a large system approach in which the number of antennas at the base station,  $N$ , and the number of mobiles in each cell,  $K$ , both grow large together, while the ratio  $\frac{K}{N}$ , which we refer to as cell loading, is held at a constant, denoted  $\beta$ .

### B. Related work

Random matrix theory has received a lot of attention in the communications literature recently [9], particularly large system results where some of the system parameters, such as number of users, the length of the signature sequence (in CDMA scenarios), or the number of transmit and receive antennas (in MIMO settings), are allowed to grow large at the same rate. Such asymptotics often produce a compact characterization of performance in the large system limit, amenable to system optimization. An attractive feature of these results is that the asymptotic expressions turn out to be good estimates for even relatively low values of the scaling parameters [10], [11], [12], [13], [14].

Most applications of large system analysis have been for the uplink (UL). Surprisingly, there has been limited work on the DL until quite recently, despite a well established UL-DL duality theory. Recently, however, a large systems analysis of regularized ZF (RZF) beamforming was carried out to characterize its limiting performance in a single cell context, allowing the optimization of the regularization parameter [15], and [16] considers ZF and RZF for correlated channel models. The present paper generalizes [15] to the multicell context, and to a wider class of beamformers, exploiting an UL-DL duality theory.

In the past year or so, there has been further work that explores the interplay between UL-DL duality, Lagrangian optimization and large systems analysis. In [17], duality between MAC and BC is used to characterize and optimize asymptotic ergodic capacity for correlated channels. A similar approach is taken in [18], [19], [20] to treat a large systems analysis of ergodic, weighted sum-rate maximization. These papers consider maximizing network utility to obtain fairness, under user scheduling, and they consider various forms of

base station cooperation (clustering). The aim is to use large system analysis to obtain numerical methods that are much more efficient than Monte-Carlo simulations; in [20] they obtain an “almost closed-form” numerical analysis tool. Linear ZF beamforming and non-ideal CSIT are considered in [19]. In [21] random matrix theory is applied to a different CBF setup than the one considered here: more particularly, they consider the problem of weighted sum of the transmit powers minimization CBF problem initially formulated in [2], and propose a strategy which also requires instantaneous local CSI and sharing channel statistics. Monte Carlo simulations are resorted to in order to claim asymptotical optimality of the results; these are however derived for a more general channel model than we use in the present paper. In [22], we provided preliminary results that we now present in more detail.

There has also been some very recent interest in using large systems analysis to succinctly answer questions concerning channel uncertainty, optimal amount of training, and required rates of feedback of channel state information. These questions are addressed for the downlink of a single cell in [23], where deterministic equivalents of SINRs for ZF and RZF are derived, under channel uncertainty, and the resulting expressions are then optimized with respect to the number of users and the number of symbols devoted to training. In [24], large systems analysis is used to optimize the number of BSs that should be cooperating on the UL, taking into account unreliable backhaul links, and the channel estimation required to measure the channels. Both papers consider much more general channel models than we do in this work.

Many duality results have been established in the context of MIMO communications. The first UL-DL duality result was obtained for the point to point MIMO channel in [25]. Another early work, [26], considered joint optimization of power and beamforming vectors for the DL of a multiple antenna cellular system employing a simple linear transmission strategy followed by single user receivers, such that the SINR at each mobile is above a target value; the proposed algorithm achieves a feasible solution for the DL if there is one and minimizes the total transmit power in the network. In [26], connections were made between this problem, and the UL power control framework of [27], and we exploit these connections in the present paper (as have many other authors). In [28], the problem of DL power minimization subject to target SINR constraints is addressed using a Lagrangian duality framework, and the transmit powers of the dual UL are found to correspond to the Lagrange coefficients associated with SINR constraints.

In the context of the capacity region of the Gaussian BC, duality results are provided in [29], [30], [31], [32]. When linear beamforming is considered, as in the present paper, several duality results have been obtained, and applied to designing iterative algorithms for DL beamforming. It was shown in [33] that, under a sum power constraint, UL and DL achievable rate regions are the same. Effective bandwidth results derived for the UL were also proven to hold in the DL. A similar approach to duality for the BC is taken in [34].

Over the past decade, a large body of research has appeared in which beamformers for the BC are derived via Lagrangian optimization techniques. Optimization formulations include

the minimization of the sum of mean-squared errors [35], [36], [37], [38], the minimization of power subject to SINR constraints [39], and the maximization of SINR's subject to power constraints [40], [39]. In [41], the problems of maximizing the sum of effective bandwidths in both the UL and DL (via duality) are considered. For the two-user case, a closed form solution may be obtained, whereas the more general problem can be formulated as a convex optimization problem and solved via an interior point method. Related problems such as SINR balancing are also considered. In [42], the Lagrangian duality approach is extended to include per-antenna power constraints. It is shown that the dual problem corresponds to an UL with uncertain noise at the receiver, and the focus is the derivation of various algorithms for solving the power minimization problem.

Some of the above-mentioned formulations are not directly convex problems; for example the DL beamforming problem to minimize power subject to SINR constraints is not a convex problem. However, it can be transformed into a convex problem, as shown in [39], which shows that strong duality holds in the original non-convex problem [42]. Thus Lagrangian methods can be applied to these non-convex problems, and this fact is exploited in the present work.

The optimization approach to beamforming provides much insight into questions of system capacity, but when it comes to practical beamforming design other considerations also come into play. For example, using an iterative algorithm to find an optimal beamformer may not be viable if channels change rapidly over time.<sup>1</sup> For this reason, many authors have focused on beamformers that can be found without iterative methods, such as the classical ZF beamformer.<sup>2</sup> Of course, rather than adapt to the instantaneous channel realizations, an alternative approach is to adapt to changing channel *statistics*. Stochastic approximation methods have a long history, and provide iterative algorithms for these scenarios. It is of interest to extend the results of the present paper to smaller (non-asymptotic) systems and such methods may prove useful.

ZF beamforming steers nulls at the other users so that they each get an interference-free version of their desired signal. However, the signal to noise penalty can be very high if the matrix to invert is ill-conditioned. RZF beamforming adds a regularization term to the ZF beamformer to provide numerical stability and better performance [44]. As noted earlier, [15] undertakes a large system analysis of RZF, exploiting how its beamformer resembles the LMMSE receiver on the UL.

A number of works have recognized that the optimal beamformer can have a RZF structure in some special cases [35], [36], [37], [40], [38], [39]. Note that for the power minimization subject to SINR constraints problem [39], the simple form of the regularized beamformer only arises when there is a great deal of symmetry in the channel model. The examples given in Section VI of [39], where such a form emerges, consist

of the diagonal case (i.e. no interference between channels) and the symmetric case, when the channel matrix has equal diagonal elements, and equal off-diagonal elements. In the present paper, channel matrices are randomly selected, so this particular symmetry condition is not satisfied. However, our model is symmetric in the average statistics of the channel matrices, which leads to the RZF beamformer resulting from the large systems limit. Nevertheless, this form of symmetry is more general than the very special one considered in [39].

In general, [39] uses iterative techniques to find numerical solutions for the optimal beamformer using conic optimization techniques. The latter are exploited in [42] to handle per-antenna power constraints, and extended further in [2], [3] to provide a duality theory for multicell systems in which there are per-BS power constraints. In [2], [3], CBf, a novel beamforming strategy in which the BS takes into account the interference it creates in adjacent cells, is proposed. [3] focuses on developing fast algorithms to find optimal beamformers, and numerical evidence is provided to show the improved performance due to coordination, relative to traditional SCP.

The RZF structure emerging from our symmetric two cell model is reminiscent of beamformers designed to optimize other criteria, such as minimum variance distortionless response (MVDR) beamformers. In particular, there is a similarity between the RZF structure and the so-called diagonal loading structure that has been used to deal with uncertainty in the noise correlation matrix [45]. In [46], a large systems analysis of MVDR beamformers with diagonal loading is undertaken where the number of antennas, and the number of observations, grow large together. Deterministic equivalents are obtained, allowing the diagonal loading factor to be optimized.

In Section III we review the optimization problems associated with the classical (non-coordinated) SCP, and the CBf strategy of [3]. We also formulate the optimization problem for MCP in which the BS antennas jointly coordinate their transmissions, as in network MIMO [47].

The main contribution of the paper is a large systems analysis of these three schemes, and optimizations based on the limiting asymptotics. In particular, we provide a clean characterization of each of the three schemes in the large systems limit in terms of effective bandwidths. In the case of SCP, we show that the RZF beamformer is asymptotically optimal in the large systems limit.<sup>3</sup> In the case of CBf, we show that the asymptotic form of the optimal beamformer provides an enhanced version of regularized ZF, one that we call generalized regularized zero-forcing (GRZF).<sup>4</sup> This beamformer is a novel contribution of the present paper.

**Notation:** Lowercase letters, boldface lowercase letters and boldface uppercase letters are used to represent scalars, vectors and matrices, respectively.  $\mathbf{x}^T$  and  $\mathbf{x}^H$  are the transpose and conjugate transpose of vector  $\mathbf{x}$ , respectively.  $\mathcal{CN}(\boldsymbol{\mu}, \mathbf{C})$  denotes a circularly symmetric complex Gaussian random vector of mean  $\boldsymbol{\mu}$  and covariance matrix  $\mathbf{C}$ . Finally  $\mathbf{S} \succeq 0$  means  $\mathbf{S}$  is a positive semidefinite (PSD) matrix.

<sup>1</sup>The classical Foschini-Miljanic[43] power control algorithm attempts to find the optimal power levels in a dynamic environment, but with beamforming there are many more parameters to be determined.

<sup>2</sup>By "iterative" we mean schemes that require iterative updates of physical parameters, such as transmit powers and beamforming vectors, not simply numerical methods that can be implemented within a single network node, e.g. in the computation of a matrix inverse.

<sup>3</sup>the structure that emerges is more general than RZF, since it also applies in regimes in which ZF is not defined.

<sup>4</sup>again, it also applies when ZF is not defined.

## II. SYSTEM MODEL

Our model, illustrated in Figure 1, has two cells, and each BS is equipped with an array of  $N$  antennas. There are  $K$  single-antenna mobiles in each cell. We assume flat fading and adopt the following notation regarding channel coefficients:

- i) the channel vector from BS  $j$  to user  $k$  in cell  $j'$  is denoted  $\mathbf{h}_{k,j',j}$ , where  $\mathbf{h}_{k,j',j} \in \mathbb{C}^{1 \times N}$ ;
- ii) the channel vector from all BSs to user  $k$  in cell  $j'$  is denoted  $\tilde{\mathbf{h}}_{k,j'}$ ; in other words,  $\tilde{\mathbf{h}}_{k,j'} = [\mathbf{h}_{k,j',1} \ \mathbf{h}_{k,j',2}]$ .

The data symbols to be transmitted to each user are assumed to be independent identically distributed (i.i.d.)  $\mathcal{CN}(0, 1)$  random variables (r.v.), and the data symbol for user  $k$  in cell  $j$  is denoted  $s_{kj}$ . Let  $\mathbf{s}_j = [s_{1j} \ \dots \ s_{Kj}]^T$  and  $\mathbf{s} = [\mathbf{s}_1^T \ \mathbf{s}_2^T]^T$ . The received signal at user  $k$  in cell  $j$  is given by

$$y_{k,j} = \sum_{j'=1}^2 \mathbf{h}_{k,j,j'} \mathbf{x}_{j'} + n_{k,j}, \quad (1)$$

where  $\mathbf{x}_{j'} \in \mathbb{C}^N$  denotes BS  $j'$ 's transmit signal, which consists of the linearly precoded symbols of the users it is serving, and which is subject to the average power constraint  $\mathbb{E}[\mathbf{x}_{j'}^H \mathbf{x}_{j'}] \leq P$  and  $n_{k,j} \sim \mathcal{CN}(0, \sigma^2)$  is the receiver noise. Receivers perform single-user detection, i.e. treat interfering signals as noise. The way the precoding vector  $\mathbf{x}_{j'}$  is generated depends on the considered scheme; See Section III.

We assume the channels between each BS and user are independent. Moreover, channels between a user and his serving BS are i.i.d.  $\mathcal{CN}(\mathbf{0}, \mathbf{I})$  whereas channels between a user and an interfering BS are i.i.d.  $\mathcal{CN}(\mathbf{0}, \epsilon \mathbf{I})$ . Thus,  $\epsilon$  controls the interference level between neighbor cells. This is a simplified model of a cellular network, yet provides useful insights [48]. It is the two cell DL version of Wyner's model [49].

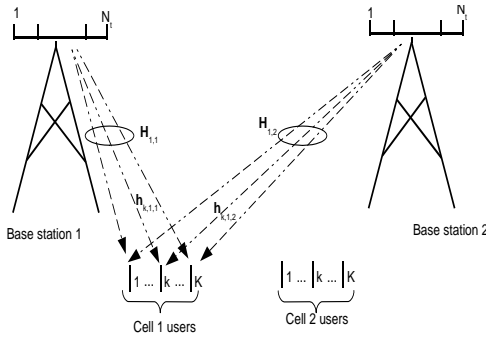


Fig. 1. System model

## III. COOPERATION AND LINEAR BEAMFORMING SCHEMES: PRIMAL PROBLEMS

We consider the problem of maximizing the network-wide minimum achievable rate for three different degrees of cooperation and coordination between the cells. The optimization to be carried out in each case is presented. The first problem is the now classic one treated in [26], the second is the cooperative scheme proposed in [3], and the third is new, although similar approaches have been proposed in several papers (e.g. [47]). In the equations below, when we consider a particular BS  $j$ , index  $\bar{j} = \text{mod}(j, 2) + 1$  corresponds to the *other* BS.

### A. Single cell processing

This is the conventional case where each BS serves its own users without worrying about the other cell. Here, we assume full re-use of time and spectrum across cells.  $\mathbf{x}_j$  is of the form:

$$\mathbf{x}_j = \sum_{k=1}^K \mathbf{w}_{kj} s_{kj} = \mathbf{W}_j \mathbf{s}_j, \quad (2)$$

where the symbol for user  $k$  in cell  $j$ ,  $s_{kj}$ , is linearly precoded by beamforming vector  $\mathbf{w}_{kj}$ .  $\mathbf{W}_j$  is the concatenation of the  $K$  precoding vectors in cell  $j$  into a  $N \times K$  matrix, the  $k$ th column being  $\mathbf{w}_{kj}$ .

In cell  $j$  the problem to be solved is the following:

$$\begin{aligned} \max_{\gamma, \mathbf{w}_{kj}, k=1, \dots, K} \quad & \gamma \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K \\ & \sum_{k=1}^K \|\mathbf{w}_{kj}\|^2 \leq P. \end{aligned} \quad (3)$$

The SINR at user  $k$  in cell  $j$  is given by

$$\text{SINR}_{k,j} = \frac{|\mathbf{h}_{k,j,j} \mathbf{w}_{kj}|^2}{\sigma_{k,j}^2 + \sum_{k'=1, k' \neq k}^K |\mathbf{h}_{k,j,j} \mathbf{w}_{k'j}|^2}, \quad (4)$$

$\sigma_{k,j}^2$  is the noise plus other-cell interference power at that user:

$$\sigma_{k,j}^2 = \sigma^2 + \sum_{k'=1}^K |\mathbf{h}_{k,j,\bar{j}} \mathbf{w}_{k'\bar{j}}|^2, \quad (5)$$

and needs to be fed back to BS  $j$ .

Solving (3) may require an iterative procedure, since beamforming at each BS influences the interference, and therefore the transmission design, at the other. We focus on finding the maximum SINR that can be met in *both* cells. This can be obtained using a bisection method. Thus, for fixed  $\gamma$  we obtain the beamforming vectors by minimizing total transmit power subject to SINR constraints on the cell's users. If the optimum is  $\leq P$  for both BSs,  $\gamma$  is attainable. We thus focus on solving:

$$\begin{aligned} \min_{\mathbf{w}_{kj}, j=1, \dots, K} \quad & \sum_{k=1}^K \|\mathbf{w}_{kj}\|^2 \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K. \end{aligned} \quad (6)$$

For  $\gamma$  achievable with unlimited transmit power, its solution will be a set of beamforming vectors that minimizes the total cell transmit power, and achieves SINR  $\gamma$ . Further, in Theorem 1, we will provide a condition that is both necessary and sufficient for the target SINR to be achievable, given unlimited power. The maximum  $\gamma$  is the target SINR for which the transmit power constraint is met with equality.

### B. Coordinated Beamforming

Here, each BS sends data to its own users only, as in SCP, but CSI is shared between the BSs so that the interference generated in other cells is taken into consideration.  $\mathbf{x}_j$  will be similar to that in (2), although the precoding design differs.

The problem formulation in (3) becomes

$$\begin{aligned} \max_{\gamma, \mathbf{w}_{k,j}, k=1, \dots, K, j=1, 2} \quad & \gamma \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K, j = 1, 2 \\ & \sum_{k=1}^K \|\mathbf{w}_{k,j}\|^2 \leq P, \quad j = 1, 2, \end{aligned} \quad (7)$$

which is a joint, two-cell optimization problem, requiring a coordinated solution. The SINR at user  $k$  in cell  $j$  is similar to (4) but can be expanded into

$$\text{SINR}_{k,j} = \frac{|\mathbf{h}_{k,j,j} \mathbf{w}_{k,j}|^2}{\sigma^2 + \sum_{j'=1}^2 \sum_{k'=1, (k',j') \neq (k,j)}^K |\mathbf{h}_{k,j,j'} \mathbf{w}_{k',j'}|^2}, \quad (8)$$

since all the channels are centrally known.

Here too the problem may be solved by a bisection method. To determine feasibility of a given  $\gamma$ , we solve (as in [3])<sup>5</sup>:

$$\begin{aligned} \min_{\phi > 0, \mathbf{w}_{k,j}, k=1, \dots, K, j=1, 2} \quad & 2P\phi \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K, j = 1, 2 \\ & \sum_{k=1}^K \mathbf{w}_{k,j}^H \mathbf{w}_{k,j} \leq \phi P, \quad \forall j = 1, 2, \end{aligned} \quad (9)$$

where  $\phi P$  upper bounds the power expended at each BS, and the aim is to use the minimum power level. Clearly, the maximum network-wide achievable SINR is the  $\gamma$  for which the optimal  $\phi = 1$ .

### C. Multicell processing

This is the case where the BSs cooperate fully: both CSI and data is available at all transmitters, who pool their antennas together to serve the users jointly. The transmitted signal  $\mathbf{x} = [\mathbf{x}_1; \mathbf{x}_2]$  will be of the form

$$\mathbf{x} = \sum_{j=1}^2 \sum_{k=1}^K \mathbf{w}_{k,j} s_{k,j} = \mathbf{W} \mathbf{s}, \quad (10)$$

where  $\mathbf{W} \in \mathbb{C}^{2N \times 2K}$  is the overall precoding matrix. The MCP optimization problem is

$$\begin{aligned} \max_{\gamma, \mathbf{w}_{k,j}, k=1, \dots, K, j=1, 2} \quad & \gamma \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K, j = 1, 2 \\ & \sum_{j'=1}^2 \sum_{k'=1}^K \|\mathbf{E}_{j,j'} \mathbf{w}_{k',j'}\|^2 \leq P, \quad j = 1, 2, \end{aligned} \quad (11)$$

where matrix  $\mathbf{E}_{j,j}$ ,  $j = 1, 2$  is diagonal and used to select the elements of each beamforming vector corresponding to BS  $j$  (i.e. its non-zero diagonal elements occupy locations  $(j-1)N+1$  to  $jN$ ). The MCP SINR at user  $k$  in cell  $j$  is

$$\text{SINR}_{k,j} = \frac{|\tilde{\mathbf{h}}_{k,j} \mathbf{w}_{k,j}|^2}{\sigma^2 + \sum_{j'=1, k'=1, (k',j') \neq (k,j)}^{2,K} |\tilde{\mathbf{h}}_{k,j} \mathbf{w}_{k',j'}|^2}. \quad (12)$$

<sup>5</sup>Note that the constant  $2P$  factor in the objective function is included as it leads to  $P$  being eliminated from the dual problem formulation.

Once again, the above problem may be solved by the bisection method. To determine feasibility of a given  $\gamma$ , we solve, as in the Cbf case, the following optimization problem:

$$\begin{aligned} \min. \quad & 2P\phi \\ \text{s.t.} \quad & \text{SINR}_{k,j} \geq \gamma, \quad k = 1, \dots, K, j = 1, 2 \\ & \sum_{j'=1}^2 \sum_{k'=1}^K \|\mathbf{E}_{j,j'} \mathbf{w}_{k',j'}\|^2 \leq \phi P, \quad j = 1, 2. \end{aligned} \quad (13)$$

## IV. SOLUTION VIA DUALITY THEORY

The above optimization problems are non-convex, and apparently nontrivial to solve. Similar problems of power control and receiver optimization on the UL are easier to solve, since each UL receive vector can be individually optimized. As discussed in Section I-B, motivated by this observation, a number of duality results have been established connecting DL optimization problems, to corresponding dual UL ones. Most relevant to us is the elegant duality theory developed by Yu and Lan [42] to handle per-antenna power constraints, exploiting earlier work on conic optimization [39].

Algorithms analogous to those proposed in [42] can be applied to solve the above problems, but even after these iterative algorithms have converged to the optimal solutions there is a great deal of complexity in the form of the optimal beamformers. Our goal is to use the duality theory to derive suboptimal (but, in the context of our simplified model, asymptotically optimal) beamformers that have relatively simple structures. In fact, generalized regularized beamformers will emerge from our analysis. Our approach will be to apply a large systems analysis to the dual, virtual UL optimization problems. Following [42], we begin by writing down the dual problems to (6), (9), and (13), respectively.

### A. Dual UL SCP problem

Let  $\lambda_{k,j}/N \geq 0$  denote the Lagrange multipliers corresponding to the SINR constraints in (6). The dual problem is

$$\begin{aligned} \max_{\lambda_{k,j} \geq 0, k=1, \dots, K} \quad & \frac{1}{N} \sum_{k=1}^K \lambda_{k,j} \sigma_{k,j}^2 \\ \text{s.t.} \quad & \mathbf{I} - \frac{\lambda_{k,j}}{N\gamma} \mathbf{h}_{k,j,j}^H \mathbf{h}_{k,j,j} + \frac{1}{N} \sum_{k' \neq k} \lambda_{k',j} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \succeq 0, \\ & k = 1, \dots, K. \end{aligned} \quad (14)$$

As explained in [42], strong duality holds despite non-convexity<sup>6</sup>, so the optimal value in the dual problem is equal to that of the primal DL beamforming problem.

The dual variables  $\lambda_{k,j}/N$  may be thought of as dual UL transmit powers, in the following dual UL power control and

<sup>6</sup>because the non-convex problem can be transformed into a convex problem using the techniques in [39]

beamforming problem [42]:

$$\begin{aligned} \min_{\mathbf{w}_{kj}, \lambda_{kj}, k=1, \dots, K} \quad & \frac{1}{N} \sum_k \lambda_{kj} \sigma_{k,j}^2 \\ \text{s.t.} \quad & \max_{\mathbf{w}_{kj}} \frac{\lambda_{kj} |\mathbf{h}_{k,j,j} \mathbf{w}_{kj}|^2}{N \mathbf{w}_{kj}^H \mathbf{w}_{kj} + \sum_{k' \neq k} \lambda_{k',j} |\mathbf{h}_{k',j,j} \mathbf{w}_{kj}|^2} \geq \gamma, \\ & k = 1, \dots, K. \end{aligned} \quad (15)$$

Indeed, (14) can be shown to have the same optimal values as (15). The optimal  $\mathbf{w}_{kj}$ , up to a scalar, are given by

$$\left( \mathbf{I} + \sum_{k' \neq k} \frac{\lambda_{k',j}}{N} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1} \mathbf{h}_{k,j,j}^H. \quad (16)$$

which we recognize to be MMSE filters for the UL problem.

To see why these problems are equivalent, note that the SINR constraints in the UL problem (15) become, after substituting for the MMSE filters:

$$\frac{\lambda_{kj}}{N} \mathbf{h}_{k,j,j} \left( \mathbf{I} + \sum_{k' \neq k} \frac{\lambda_{k',j}}{N} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1} \mathbf{h}_{k,j,j}^H \geq \gamma. \quad (17)$$

The optimal  $\lambda_{kj}$ s are the unique solution to the fixed point equation:

$$\lambda_{kj} = \frac{\gamma N}{\mathbf{h}_{k,j,j} \left( \mathbf{I} + \sum_{k' \neq k} \frac{\lambda_{k',j}}{N} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1} \mathbf{h}_{k,j,j}^H}. \quad (18)$$

Since  $\gamma$  is achieved exactly, the SINR constraints in (15) are achieved with equality. It follows that if the minimization in (15) is changed to a maximization, and the SINR inequalities are reversed, then the same solution will be found. It is shown in [42] that in so doing, one obtains (14).

The right hand side of (18) is an *interference function* of the dual UL powers, in the sense of Yates' framework on UL power control [27]. Thus, iterative approaches to power control converge to the optimal dual variables. Indeed, this was the original approach to the multicell beamforming problem taken in [26], one of the first papers to exploit an UL-DL duality. More recently, more efficient approaches to solving the dual problem have been proposed [42], but these still require several iterations before the algorithm converges, which could be problematic. The large systems analysis in Section V allows us to derive simple yet asymptotically optimal beamformers.

From the Karush-Kuhn-Tucker (KKT) conditions, one can show that the optimal  $\mathbf{w}_{kj}$  for the primal are the same, up to a scaling factor, as  $\hat{\mathbf{w}}_{kj}$ . Thus, they can be written as

$$\mathbf{w}_{kj} = \sqrt{\frac{p_{kj}}{N}} \frac{\hat{\mathbf{w}}_{kj}}{\|\hat{\mathbf{w}}_{kj}\|}, \quad (19)$$

where  $\frac{p_{kj}}{N}$  is the transmit power allocated to beamforming vector  $\mathbf{w}_{kj}$  on the DL. From the DL SINR constraints, for  $k = 1, \dots, K$ ,

$$\frac{p_{kj}}{N\gamma} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{kj}|^2}{\|\hat{\mathbf{w}}_{kj}\|^2} - \sum_{k' \neq k} \frac{p_{k',j}}{N} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2} = \sigma_{k,j}^2, \quad (20)$$

and the  $(p_{kj})_{k=1}^K$  can be determined from this set of equations. We undertake a large system analysis of this scheme in Section V, and present the results in Theorem 1.

As a final remark, we note that the UL-DL duality described in this section is not the same as the celebrated UL-DL duality used to characterize the capacity region of the MIMO BC. In the present section (and indeed throughout this paper), we do not allow so-called dirty paper coding (DPC), which is the coding technique that achieves the capacity of the MIMO BC. The original notion of UL-DL duality was a correspondence between points in the capacity region of the MIMO BC with points in a dual UL multiple access channel with suitably chosen noise covariances. However, it has been shown that the optimization approach developed in [42] to obtain the optimal beamforming vectors for the DL power minimization problem (as reviewed in the present section) can be extended to allow DPC, and hence to obtain the MIMO BC capacity results, including the setup with per antenna power constraints (see [42]).

### B. Dual UL CBf problem

Letting  $\lambda_{kj}/N \geq 0$  denote the Lagrange multipliers corresponding to the SINR constraints and  $\mu_j$  those corresponding to the power constraints in (9), its Lagrangian dual problem is

$$\begin{aligned} \max_{\lambda_{kj} \geq 0, \mu_j \geq 0} \quad & \sum_{j,k} \frac{\lambda_{kj}}{N} \sigma^2 \\ \text{s.t.} \quad & \mu_j \mathbf{I} - \frac{\lambda_{kj}}{\gamma N} \mathbf{h}_{k,j,j}^H \mathbf{h}_{k,j,j} \\ & + \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \succeq 0, \\ & k = 1, \dots, K, j = 1, 2 \\ & \sum_{j=1}^2 (1 - \mu_j) P = 0. \end{aligned} \quad (21)$$

As for the SCP case, this problem can be shown to be equivalent to a dual UL problem with uncertain noise (cf. [42]), or equivalently

$$\begin{aligned} \max_{\mu_j \geq 0} \min_{\lambda_{kj} \geq 0, \mathbf{w}_{kj}} \quad & \sum_{k,j} \frac{\lambda_{kj}}{N} \sigma^2 \\ \text{s.t.} \quad & \frac{\lambda_{kj} |\hat{\mathbf{w}}_{kj} \mathbf{h}_{k,j,j}|^2}{\hat{\mathbf{w}}_{kj}^H \left[ N \mu_j \mathbf{I} + \sum_{(k',j') \neq (k,j)} \lambda_{k',j'} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \right] \hat{\mathbf{w}}_{kj}} \geq \gamma, \\ & k = 1, \dots, K, j = 1, 2 \\ & \sum_{j=1}^2 (1 - \mu_j) = 0. \end{aligned} \quad (22)$$

The Lagrange multipliers  $\mu_1, \mu_2$  can be interpreted as noise levels at the two BSs, respectively, on the virtual UL [42].

For any choice of dual variables,  $(\mu_1, \mu_2)$ ,  $(\frac{1}{N} \lambda_{kj})_{k=1, j=1}^{K, 2}$ , the SINR achieved on the virtual UL can be shown to be:

$$\frac{\lambda_{kj}}{N} \mathbf{h}_{k,j,j} \left[ \mu_j \mathbf{I} + \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \right]^{-1} \mathbf{h}_{k,j,j}^H. \quad (23)$$

Using the optimal dual variables, all (23) equal  $\gamma$ , which allows us to write down the optimal  $(\lambda_{kj})_{k=1,j=1}^{K,2}$  as the unique solution to a fixed point equation:

$$\lambda_{kj} = \frac{\gamma N}{\mathbf{h}_{k,j,j} \left[ \mu_j \mathbf{I} + \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \right]^{-1} \mathbf{h}_{k,j,j}^H}. \quad (24)$$

Once  $(\mu_1, \mu_2)$  are determined, this equation provides an implicit solution to the dual problem, in that standard iterative methods can be applied to obtain the dual UL powers. It remains to find  $\mu_1$  and  $\mu_2$ . This is addressed in Section V.

The optimal beamforming vectors on the dual UL,  $\hat{\mathbf{w}}_{kj}$ , assuming feasibility, are (up to a scaling factor)

$$\left[ \mu_j \mathbf{I} + \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \right]^{-1} \mathbf{h}_{k,j,j}^H; \quad (25)$$

the optimal beamforming vectors on the DL are of the form  $\mathbf{w}_{kj} = \sqrt{\frac{p_{kj}}{N}} \frac{\hat{\mathbf{w}}_{kj}}{\|\hat{\mathbf{w}}_{kj}\|}$ , where  $\frac{p_{kj}}{N}$  is the power allocated to the beamforming vector  $\mathbf{w}_{kj}$ . From DL SINR constraints, for  $j = 1, 2, k = 1, \dots, K$ ,

$$\frac{p_{kj}}{N\gamma} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{kj}|^2}{\|\hat{\mathbf{w}}_{kj}\|^2} - \sum_{j',k', (k',j') \neq (k,j)} \frac{p_{k',j'}}{N} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k',j'}|^2}{\|\hat{\mathbf{w}}_{k',j'}\|^2} = \sigma^2. \quad (26)$$

### C. Dual UL MCP problem

With  $\frac{\lambda_{kj}}{N}$ 's and  $\mu_j$ 's defined as in (IV-B), the Lagrangian dual to problem (13) is equivalent to

$$\max_{\mu_j \geq 0} \min_{\lambda_{kj} \geq 0, \hat{\mathbf{w}}_{kj}} \sum_{j=1}^2 \sum_{k=1}^K \frac{\lambda_{kj}}{N} \sigma^2 \quad (27)$$

$$\text{s.t. } \frac{\frac{\lambda_{kj}}{N} |\tilde{\mathbf{h}}_{k,j} \hat{\mathbf{w}}_{kj}|^2}{\hat{\mathbf{w}}_{kj}^H \left[ \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{M} \right] \hat{\mathbf{w}}_{kj}} \geq \gamma \quad (28)$$

$$\sum_{j=1}^2 (1 - \mu_j) = 0. \quad (29)$$

where, to simplify the equations,  $\mathbf{M} = \sum_{j'=1}^2 \mu_{j'} \mathbf{E}_{j'}$  is used.

The optimal  $\hat{\mathbf{w}}_{kj}$  is

$$\left[ \frac{1}{N} \sum_{(k',j') \neq (k,j)} \lambda_{k',j'} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{M} \right]^{-1} \tilde{\mathbf{h}}_{k,j}^H. \quad (30)$$

Plugging this into the inequality in (28), we obtain

$$\frac{\lambda_{kj}}{N} \tilde{\mathbf{h}}_{k,j} \left[ \sum_{(k',j') \neq (k,j)} \frac{\lambda_{k',j'}}{N} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{M} \right]^{-1} \tilde{\mathbf{h}}_{k,j}^H \geq \gamma. \quad (31)$$

Define  $\check{\mathbf{H}}_{k,j}$  as

$$\check{\mathbf{H}}_{k,j} = \begin{bmatrix} \check{\mathbf{h}}_{1,j,j} \\ \vdots \\ \check{\mathbf{h}}_{k-1,j,j} \\ \check{\mathbf{h}}_{k+1,j,j} \\ \vdots \\ \check{\mathbf{h}}_{K,j,j} \\ \check{\mathbf{h}}_{1,\bar{j},j} \\ \vdots \\ \check{\mathbf{h}}_{K,\bar{j},j} \end{bmatrix}, \quad (32)$$

where  $\check{\mathbf{h}}_{k,j} = \tilde{\mathbf{h}}_{k,j} \mathbf{M}^{-1/2}$ , and let  $\mathbf{L}_{k,j}$  is the diagonal matrix whose diagonal entries vector is given by

$$[\lambda_{1j}; \dots; \lambda_{(k-1)j}; \lambda_{(k+1)j}; \dots; \lambda_{Kj}; \lambda_{1\bar{j}}; \dots; \lambda_{K\bar{j}}]. \quad (33)$$

Using this notation, for any choice of dual variables,  $(\mu_1, \mu_2)$ ,  $\left(\frac{\lambda_{kj}}{N}\right)_{k=1,j=1}^{K,2}$ , the SINR achieved on the virtual UL (i.e. the left-hand side of Eq. (31)) can be rewritten as

$$\frac{\lambda_{kj}}{N} \check{\mathbf{h}}_{k,j} \left[ \frac{1}{N} \check{\mathbf{H}}_{k,j}^H \mathbf{L}_{k,j} \check{\mathbf{H}}_{k,j} + \mathbf{I} \right]^{-1} \check{\mathbf{h}}_{k,j}^H. \quad (34)$$

With the *optimal* dual variables, all (34) equal  $\gamma$ , which allows us to write down the optimal  $(\lambda_{kj})_{k=1,j=1}^{K,2}$  as the solution of a fixed point equation:

$$\lambda_{kj} = \gamma N \left( \check{\mathbf{h}}_{k,j} \left[ \frac{1}{N} \check{\mathbf{H}}_{k,j}^H \mathbf{L}_{k,j} \check{\mathbf{H}}_{k,j} + \mathbf{I} \right]^{-1} \check{\mathbf{h}}_{k,j}^H \right)^{-1}. \quad (35)$$

As for CBf, once  $\mu_1, \mu_2$  are obtained, the optimal  $(\lambda_{kj})_{k=1,j=1}^{K,2}$  are implicitly given as the unique solution of (35). As the duality gap is zero for a feasible primal, at the optimum,

$$2P\phi = \frac{1}{N} \sum_{j=1}^2 \sum_{k=1}^K \lambda_{kj} \sigma^2. \quad (36)$$

Moreover, the beamforming vectors in the original problem and in the dual problem are related as follows:

$$\mathbf{w}_{kj} = \sqrt{\frac{p_{kj}}{N}} \frac{\hat{\mathbf{w}}_{kj}}{\|\hat{\mathbf{w}}_{kj}\|}. \quad (37)$$

Plugging these into the DL SINR constraints provides the solution for the  $p_{kj}$ :

$$\frac{p_{kj}}{N\gamma} \frac{|\tilde{\mathbf{h}}_{k,j} \hat{\mathbf{w}}_{kj}|^2}{\|\hat{\mathbf{w}}_{kj}\|^2} - \sum_{(k',j') \neq (k,j)} \frac{p_{k',j'}}{N} \frac{|\tilde{\mathbf{h}}_{k,j} \hat{\mathbf{w}}_{k',j'}|^2}{\|\hat{\mathbf{w}}_{k',j'}\|^2} = \sigma^2. \quad (38)$$

## V. LARGE SYSTEM RESULTS

We proceed to a large system analysis of the solutions to the above optimization problems, in the limit as  $N, K$  grow large, keeping the cell loading ratio  $\beta = \frac{K}{N}$  fixed. The basic idea is to apply large systems analysis techniques to the dual UL problems, guided by the analysis of similar UL problems in the literature [10], [50]. For example, the SINR expression in (17) is typically analyzed by considering the limit of the empirical distribution of the eigenvalues of the matrix  $\left( \mathbf{I} + \frac{1}{N} \sum_{k' \neq k} \lambda_{k',j} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1}$  as  $N, K \rightarrow \infty$ ,

with  $K/N$  held at  $\beta$ . Once this is characterized, a law of large numbers for the trace of this matrix can be obtained, and from that, the limit of the SINR expression in (17).

In the present paper, extra technical difficulties arise. To directly apply the known theorems about large random matrices, we need to assume that the dual UL powers  $\lambda_{kj}$  are *independent* of the channel matrix parameters, and are chosen *independently* of each other from some fixed distribution. However, being solutions of the UL dual problem, they form a collection of *dependent* random variables which, moreover, depend on the channel matrix parameters. Our approach to this problem is to provide lower and upper bounds to the SINR's using *deterministic* dual UL powers. We obtain these bounds by exploiting the underlying monotonicity structure of the UL power control problem [51], [27], [26]. With deterministically chosen UL powers, we can perform a large system analysis, and then provide a sandwich type argument to show that the optimal dual UL powers must converge to deterministic values. Since this part of the paper is somewhat technical, we relegate it to the appendices, and focus on the results of the analysis.

#### A. Asymptotically Optimal Beamformers

**Theorem 1** (Asymptotically optimal beamforming for SCP). *Assume  $\beta \left( \frac{\gamma}{1+\gamma} + \epsilon\gamma \right) < 1$ . Then, asymptotically, SINR  $\gamma$  is achievable at each mobile terminal in the limit as  $N \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta > 0$ . In this case, the empirical distribution of the (normalized) UL dual power levels (i.e. the  $\lambda_{kj}$ s) converges weakly to the constant  $\bar{\lambda}$  given by*

$$\bar{\lambda} = \frac{\gamma}{1 - \beta \frac{\gamma}{1+\gamma}}, \quad (39)$$

*the empirical distribution of the (normalized) DL power per users (i.e. the  $p_{kj}$ s) converges weakly to the constant*

$$\bar{p} = \frac{\sigma^2 \gamma}{1 - \beta \left( \frac{\gamma}{1+\gamma} + \epsilon\gamma \right)}. \quad (40)$$

*The per BS power converges to  $\bar{P} = \beta \bar{p}$ . The asymptotically optimal form of the DL beamformer for user  $k$  in cell  $j$  is*

$$\mathbf{w}_{kj}^{SCP} = \sqrt{\frac{\bar{p}}{N}} \frac{\hat{\mathbf{w}}_{kj}^{SCP}}{\|\hat{\mathbf{w}}_{kj}^{SCP}\|}, \quad (41)$$

$$\text{where } \hat{\mathbf{w}}_{kj}^{SCP} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{k' \neq k} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1} \mathbf{h}_{k,j,j}^H.$$

*Finally, the asymptotic SINR,  $\gamma$ , is related to the other variables via the fixed point equations*

$$\gamma = \frac{1}{\frac{1}{\bar{\lambda}} + \frac{\beta}{1+\gamma}} = \frac{1}{\frac{\sigma^2}{\bar{p}} + \epsilon\beta + \frac{\beta}{1+\gamma}}. \quad (42)$$

*Conversely, if  $\beta \left( \frac{\gamma}{1+\gamma} + \epsilon\gamma \right) > 1$ , then, asymptotically, the SINR target  $\gamma$  is not achievable under the SCP strategy.*

*Proof:* See Appendix D. ■

**Corollary 1** (SCP). *Subject to the per BS power constraint  $P$ , the maximum asymptotic network-wide achievable SINR for a*

*given cell loading factor  $\beta$  is the unique positive solution to the following fixed point equation:*

$$\gamma_{SCP}^* = \frac{1}{\beta} \frac{1}{\frac{\sigma^2}{P} + \epsilon + \frac{1}{1+\gamma_{SCP}^*}}, \quad (43)$$

*which has the explicit solution  $\gamma_{SCP}^*$  equal to*

$$\frac{-\left(\frac{\sigma^2}{P} + \epsilon - \frac{1}{\beta} + 1\right) + \sqrt{\left(\frac{\sigma^2}{P} + \epsilon - \frac{1}{\beta} + 1\right)^2 + 4\frac{\frac{\sigma^2}{P} + \epsilon}{\beta}}}{2\left(\frac{\sigma^2}{P} + \epsilon\right)}.$$

Theorem 1 is interesting in that it relates the solution to the optimization problem (6) to a notion of *regularized zero-forcing* proposed in [44] as a *practical* approach that is as simple to implement as ZF, yet with better performance. It was studied in the asymptotic regime in [15]. The beamformers defined in (41) asymptotically lead to a precoding matrix

$$\mathbf{W}_j^{SCP} = c_j \left[ \mathbf{I} + \frac{\bar{\lambda}}{N} \mathbf{H}_{j,j}^H \mathbf{H}_{j,j} \right]^{-1} \mathbf{H}_{j,j}^H, \quad (44)$$

where  $c_j$  ensures the power constraint at  $BS_j$  is met with equality, and  $\mathbf{H}_{j,j}$  is the concatenation of channels between cell  $j$  users and their serving BS. We should add that the optimal beamformer always exists, even when the ZF beamformer does not, so technically we should only refer to it as RZF in those scenarios where the ZF beamformer exists.

Another interesting observation is that Theorem 1 provides a condition that is both necessary and sufficient for the target SINR,  $\gamma$ , to be achievable. We can interpret  $\frac{\gamma}{1+\gamma}$  as the *effective bandwidth* of a user in cell  $j$ , and  $\epsilon\gamma$  as the effective bandwidth of an interferer in cell  $\bar{j}$ . Effective bandwidths provide a simple metric by which different beamforming schemes can be compared, as shown in the next two theorems.

**Theorem 2** (Asymptotically optimal beamforming for CBf). *Assume  $\beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right) < 1$ . Then, asymptotically, SINR  $\gamma$  is achievable at each mobile terminal, in the limit as  $N \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta > 0$ . In this case, the empirical distribution of the (normalized) UL dual power levels (i.e. the  $\lambda_{kj}$ s) converges weakly to the constant  $\bar{\lambda}$ , given by*

$$\bar{\lambda} = \frac{\gamma}{1 - \beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right)}, \quad (45)$$

*the empirical distribution of the (normalized) DL power per users (i.e. the  $p_{kj}$ s) converges weakly to the constant*

$$\bar{p} = \bar{\lambda} \sigma^2 = \frac{\sigma^2 \gamma}{1 - \beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right)}. \quad (46)$$

*The per BS power converges to  $\bar{P} = \beta \bar{p}$ . The asymptotically optimal form of the DL beamformer for user  $k$  in cell  $j$  is*

$$\mathbf{w}_{kj}^{Coord} = \sqrt{\frac{\bar{p}}{N}} \frac{\hat{\mathbf{w}}_{kj}^{Coord}}{\|\hat{\mathbf{w}}_{kj}^{Coord}\|}, \quad (47)$$

$$\text{where } \hat{\mathbf{w}}_{kj}^{Coord} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{(k', \bar{j}) \neq (k, j)} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} \right)^{-1} \mathbf{h}_{k,j,j}^H.$$



Finally, the asymptotic SINR,  $\gamma$ , is related to the other variables via

$$\gamma = \frac{1}{\frac{1}{\bar{\lambda}} + \frac{\epsilon\beta}{1+\epsilon\gamma} + \frac{\beta}{1+\gamma}} = \frac{1}{\frac{\sigma^2}{\bar{p}} + \frac{\epsilon\beta}{1+\epsilon\gamma} + \frac{\beta}{1+\gamma}}. \quad (48)$$

Conversely, if  $\beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right) > 1$ , then asymptotically the SINR target  $\gamma$  is not achievable under the coordinated beamforming strategy.

*Proof:* See Appendix E. ■

**Corollary 2 (CBf).** Subject to per BS power constraint  $P$ , the maximum asymptotic network-wide achievable SINR for a given cell loading factor  $\beta$  is the unique positive solution to the following fixed point equation:

$$\gamma_{CBf}^* = \frac{1}{\beta} \frac{1}{\frac{\sigma^2}{P} + \frac{\epsilon}{1+\epsilon\gamma_{CBf}^*} + \frac{1}{1+\gamma_{CBf}^*}}. \quad (49)$$

In other words,  $\gamma_{CBf}^*$  is the root of a cubic equation.

Theorem 2 is interesting in that it provides a *novel* form of RZF beamformer.<sup>7</sup> This beamformer is not a direct regularization of the standard ZF beamformer in a single cell. Rather, it is a regularization of a beamformer that zero forces the interference it creates at users in the other cell as well as its own; in other words, it transmits to a *subset* of the users that it is zero forcing. We call this a generalized RZF beamformer, and it is a novel contribution of the present paper. It can be used in a finite system, where it is suboptimal, but relatively straightforward to implement. Theorem 2 shows that it is asymptotically optimal in the class of coordinated beamforming strategies. Note also the clean characterization of effective bandwidth for this beamformer, and that it provides a significant reduction in the effective bandwidth of the other-cell users, compared to SCP, when  $\epsilon$  is non-negligible.

**Theorem 3** (Asymptotically optimal beamforming for MCP). Assume  $\beta \frac{\gamma}{1+\gamma} < 1$ . Then, asymptotically, SINR  $\gamma$  is achievable at each mobile terminal, in the limit as  $N \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta > 0$ . In this case, the empirical distribution of the (normalized) UL dual power levels (i.e. the  $\lambda_{kj}$ s) converges weakly to the constant  $\bar{\lambda}$  given by

$$\bar{\lambda} = \frac{1}{1+\epsilon} \frac{\gamma}{\left(1 - \beta \frac{\gamma}{1+\gamma}\right)}, \quad (50)$$

the empirical distribution of the (normalized) DL power per user (i.e. the  $p_{kj}$ s) converges weakly to the constant

$$\bar{p} = \bar{\lambda}\sigma^2 = \frac{1}{1+\epsilon} \frac{\sigma^2\gamma}{\left(1 - \beta \frac{\gamma}{1+\gamma}\right)}. \quad (51)$$

The per BS power converges to  $P = \beta\bar{p}$ . The asymptotically optimal form of the DL beamformer for user  $k$  in cell  $j$  is

$$\mathbf{w}_{kj}^{MCP} = \sqrt{\frac{\bar{p}}{N}} \frac{\hat{\mathbf{w}}_{kj}^{MCP}}{\|\hat{\mathbf{w}}_{kj}^{MCP}\|}, \quad (52)$$

<sup>7</sup>in the scenarios where the ZF beamformer exists.

$$\text{where } \hat{\mathbf{w}}_{kj}^{MCP} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{(k',j') \neq (k,j)} \tilde{\mathbf{h}}_{k',j}^H \tilde{\mathbf{h}}_{k',j} \right)^{-1} \tilde{\mathbf{h}}_{k,j}^H. \quad (53)$$

Finally, the asymptotic SINR,  $\gamma$ , is related to the other variables via

$$\gamma = \frac{1}{\frac{1}{(1+\epsilon)\bar{\lambda}} + \frac{\beta}{1+\gamma}} = \frac{1}{\frac{\sigma^2}{\bar{p}(1+\epsilon)} + \frac{\beta}{1+\gamma}}. \quad (54)$$

Conversely, if  $\beta \frac{\gamma}{1+\gamma} > 1$ , then asymptotically the SINR target  $\gamma$  is not achievable under the MCP beamforming strategy.

*Proof:* See Appendix F. ■

**Corollary 3 (MCP).** Subject to per BS power constraint  $P$ , the maximum asymptotic network-wide achievable SINR for a given cell loading factor  $\beta$  is the unique positive solution to the following fixed point equation:

$$\gamma_{MCP}^* = \frac{1}{\beta} \frac{1}{\frac{\sigma^2}{(1+\epsilon)P} + \frac{1}{1+\gamma_{MCP}^*}}. \quad (55)$$

In other words,  $\gamma_{MCP}^*$  is equal to

$$\frac{-\left(\frac{\sigma^2}{(1+\epsilon)P} - \frac{1}{\beta} + 1\right) + \sqrt{\left(\frac{\sigma^2}{(1+\epsilon)P} - \frac{1}{\beta} + 1\right)^2 + 4\frac{\frac{\sigma^2}{(1+\epsilon)P}}{\beta}}}{2\left(\frac{\sigma^2}{(1+\epsilon)P}\right)}.$$

Although MCP is a complex strategy, in that the cooperation between BSs is much greater, the reward is better performance than that attainable in a single isolated cell with no intercell interference. Note that the power levels in (51) are *less* than what they would be in a single isolated cell. The effective bandwidth of each user is the same as for a single isolated cell, under SCP, but the power consumption is reduced by the factor  $(1+\epsilon)$ , which corresponds to a power gain from having both BSs involved in the beamforming, instead of just one.

## B. Effective interference

The optimal SINR expressions in (42), (48), (54) are striking in how they capture the effect of interference for the three different beamformers. Indeed, they supply a simple “effective interference” characterization, which can be used to directly check if a particular target SINR can be achieved.

It is natural to try and compare the schemes directly using the limiting SINR expressions. This is accomplished in the following theorem, where  $SNR$  denotes  $\frac{P}{\sigma^2}$ .

**Theorem 4.** Let  $\gamma_{SCP}^*, \gamma_{CBf}^*, \gamma_{MCP}^*$  denote the SINRs under SCP, CBf, and MCP, respectively. Then

$$\gamma_{SCP}^* < \gamma_{CBf}^* < \gamma_{MCP}^*. \quad (56)$$

At signal to noise ratio  $SNR$  and interference level  $\epsilon$ , denote the effective interference at target SINR  $\gamma$  by

$$I_{\text{eff}}(SNR, \epsilon, \gamma) = \begin{cases} \beta \left(1 + \frac{SNR}{1+\gamma} + \epsilon SNR\right) & \text{SCP} \\ \beta \left(1 + \frac{SNR}{1+\gamma} + \frac{\epsilon SNR}{1+\epsilon\gamma}\right) & \text{CBf} \\ \beta \left(1 + \frac{SNR}{1+\gamma} + \frac{\epsilon SNR}{1+\gamma}\right) & \text{MCP} \end{cases}$$

Then the feasibility of  $\gamma$  in the case of SCP, or CBf, is equivalent to satisfaction of the inequality  $\frac{SNR}{I_{\text{eff}}(SNR, \epsilon, \gamma)} > \gamma$ , and in the MCP case, it is equivalent to  $\frac{(1+\epsilon)SNR}{I_{\text{eff}}(SNR, \epsilon, \gamma)} > \gamma$ .

*Proof:* Follows closely that of Proposition 3.2 in [10]. ■

We note here the close parallel with the effective interference arising in the large system analysis of linear UL multiuser receivers [10]. This is due to the underlying UL-DL duality.

### C. Asymptotically optimal cell loading

The above theorems characterize the optimal SINR for fixed cell loading  $\beta$  under SCP, CBf, and MCP, respectively. Our next step is to characterize this optimum loading: this determines the optimal number of users to serve given a number of antennas at the BS. This corresponds to finding the  $\beta$  that maximizes the normalized (by the number of antennas) rate per cell  $r$ , i.e. the optimizing  $\beta$  that solves the following problem:

$$\text{maximize}_{\beta} r = \beta \log(1 + \gamma^*) \quad (57)$$

with  $\gamma^*$  characterized by the appropriate fixed point equation (cf. Eqs (43), (49) and (55)).

**Proposition 1** (Characterization of the optimum  $\beta$  for SCP). *If*

$$\epsilon + \frac{\sigma^2}{P} \geq 1 \quad (58)$$

*then  $r(\beta)$  is an increasing function. Otherwise, the optimum occurs at a finite  $\beta^*$  which may be found by a line search.*

*Proof:* Refer to Appendix G. ■

**Proposition 2** (Optimal cell loading for CBf). *If  $\frac{\sigma^2}{P} + \epsilon - 2\epsilon^2 - 1 \geq 0$  then  $r(\beta)$  is an increasing function. Otherwise, there is a finite value of  $\beta$  at which  $r$  is maximized.*

*Proof:* Refer to Appendix H. ■

**Proposition 3** (Characterization of the optimum  $\beta$  for MCP). *If*

$$\frac{\sigma^2}{P} \geq (1 + \epsilon) \quad (59)$$

*then  $r(\beta)$  is an increasing function. Otherwise, the optimum occurs at a finite  $\beta^*$  which may be found by a line search.*

*Proof:* Comparing (55) and (43), we see that the former is the same as the latter with  $\epsilon + \frac{\sigma^2}{P} \leftarrow \frac{\sigma^2}{P(1+\epsilon)}$ . Performing this substitution in (58) yields the result. ■

The above results define for each scheme a noise-limited region, in which cell loading can be increased indefinitely; however, this leads to ever decreasing rates per user, not to mention that more user channels would have to be learned.

## VI. PERFORMANCE RESULTS

How do these schemes compare with each other and with other approaches from the literature? Clearly, CBf requires more CSI than SCP, and MCP involves much more BS

cooperation, so it is not surprising that the SINRs are ordered as in (56). In this section, we obtain numerical results to provide a quantitative comparison between these schemes in different scenarios. Throughout this paper, we have assumed full re-use of time and spectrum across cells; however, intercell interference can be avoided altogether by applying the classic principle of re-use partitioning: we thus also consider in our simulations the time division (TD) scheme in which each BS is given a separate time-slot, which we shall also call “1/2-reuse”. We also consider two forms of pure ZF in the context of single cell processing. SCP-ZF zero forces the same-cell interference, with the BS oblivious to other-cell interference. Generalized zero-forcing (GZF) is when the BSs independently zero force the interference in the two-cell system. Finally, in the MCP setting, the two BSs can jointly zero-force all the interference in the two-cell system, and we denote this case by “MCP-ZF”.

### A. When is half re-use SCP better than CBf?

Let  $\beta_{TD}$  be the cell loading in the half reuse scheme in which each BS transmits half the time. To compare with coordinated beamforming (a full re-use scheme), let  $\beta := \beta_{CBf} := \beta_{TD}/2$ . Then the rate for the TD scheme is

$$r_{TD}(\beta) = \beta \log(1 + \gamma_{TD}^*(\beta)), \quad (60)$$

$$\text{with } \gamma_{TD}^*(\beta) = \frac{1}{2\beta} \frac{1}{\frac{\sigma^2}{2P} + \frac{1}{1+\gamma_{TD}^*(\beta)}} = \frac{1}{\beta} \frac{1}{\frac{\sigma^2}{P} + \frac{2}{1+\gamma_{TD}^*(\beta)}}. \quad (61)$$

The rate using coordinated beamforming is given by

$$r_{CBf}(\beta) = \beta \log(1 + \gamma_{CBf}^*(\beta)) \quad (62)$$

$$\text{with } \gamma_{CBf}^* = \frac{1}{\beta} \frac{1}{\frac{\sigma^2}{P} + \frac{1}{1+\gamma_{CBf}^*(\beta)} + \frac{\epsilon}{1+\epsilon\gamma_{CBf}^*(\beta)}}. \quad (63)$$

Thus,  $r_{CBf}(\beta) > r_{TD}(\beta)$  if  $\epsilon < 1$ , and  $r_{CBf}(\beta) < r_{TD}(\beta)$  if  $\epsilon > 1$ . It follows that coordinated beamforming is only useful when  $\epsilon < 1$ ; otherwise, it is better to partition the cells with a reuse factor of 1/2. Of course, if BS association is performed properly,  $\epsilon$  should be less than 1.

### B. Numerical results

Figures 2-4 compare the different schemes by varying the cell loading  $\beta$ . We notice that when  $\epsilon$  is small, CBf gains little over SCP, but offers significant gains compared to pure ZF or 1/2-reuse. When  $\epsilon$  is small, SCP-ZF is superior to GZF, as expected. When  $\epsilon$  is large, but  $< 1$  (e.g.  $\epsilon = 0.8$ ), CBf gains significantly over SCP, but does not gain much over 1/2-reuse, or to GZF (when the loading is perfectly optimized). Note that in this case, the relevant comparison is with GZF.

When  $\epsilon < 1$ , CBf is always better, for appropriately selected  $\beta$ , than SCP, SCP-ZF, and GZF. If one were to insist on using pure ZF, and could choose the appropriate ZF scheme, and the exact optimal loading for that scheme, then it can get reasonably close to the performance of the optimized

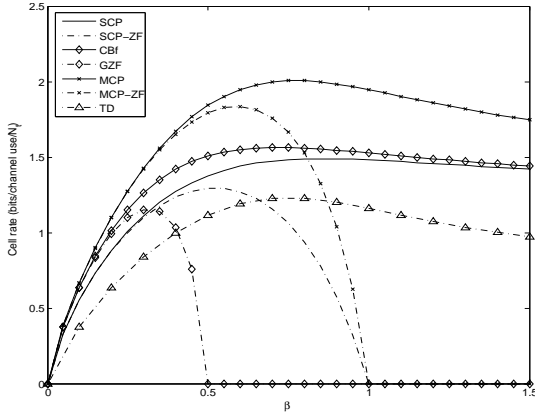


Fig. 2. Effect of cell loading  $\beta$  on rate achieved for  $\text{SNR} = 10\text{dB}$ ,  $\epsilon = .1$

CBf. Similarly, if one could select between SCP or 1/2-reuse, the performance can be quite close to that of CBf. The advantage of the latter is that it is universally good, if no joint transmission is allowed, across all levels of inter-cell interference. Compared to the ZF schemes, it performs better across a wider range of cell loadings. In large networks, it avoids the intractable frequency planning problem associated with fractional re-use schemes. When  $\epsilon > 1$ , MCP offers very significant gains over the 1/2-reuse scheme (not depicted).

In the two-cell model, MCP offers the most gain when  $\epsilon$  is large. Even when  $\epsilon$  is small, as in Figure 2, the gains over CBf, 1/2-reuse, and single cell ZF schemes, respectively, are significant, and in Figure 4 they are higher still, because  $\epsilon$  is larger in that case. Unlike the other schemes, MCP improves with increasing  $\epsilon$ , but it requires significant cooperation between BSs, including full data sharing, whose cost in terms of backhaul capacity is not accounted for here.

Finally, we investigate the applicability of the asymptotic results to a finite system. In a first step, for  $K = 3$ ,  $N = 4$ ,  $\frac{P}{\sigma^2} = 10$  and  $\epsilon$  taking values in  $[.01; .1; .5; .8; 1]$ , we solve the optimization problems described in Section III for different independent samples of the channel and obtain the corresponding average rates. Even for such a small number of antennas, the large system analysis (LSA) results provide quite a good approximation. The results are shown in Figure 5.

The optimizations in Section III are time-consuming, particularly for the SCP case, which requires iterations between the optimization at the two transmitters until convergence. Thus, we consider applying the asymptotically optimal beamforming vectors from Section V (slightly modified so as not to break the per transmitter power constraint for the MCP scheme) in the finite system case. The results are shown in Figure 6.

## VII. CONCLUSIONS

This paper has provided an asymptotic analysis of a two cell interfering network in which the number of antennas at the BSs, and that of users per cell, both grow large together. Schemes that balance rates across users in the system were compared for three levels of cooperation, namely single cell processing, coordinated beamforming and multicell processing. MCP offers significant rate gains, if we can accommodate the additional coordination and communication between BSs,

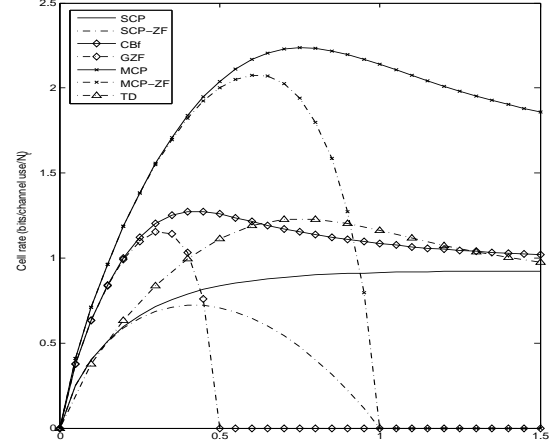


Fig. 3. Effect of cell loading on rate achieved for  $\text{SNR} = 10\text{dB}$ ,  $\epsilon = .5$

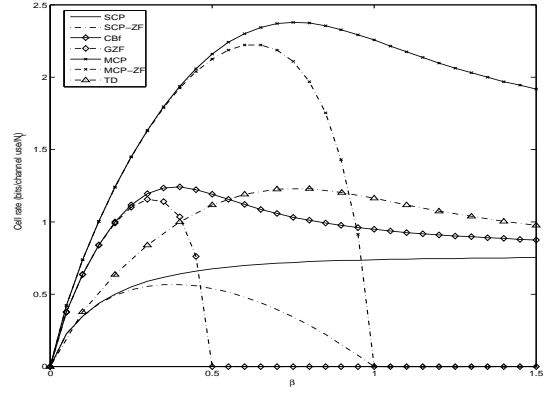


Fig. 4. Effect of cell loading on rate achieved for  $\text{SNR} = 10\text{dB}$ ,  $\epsilon = .8$

not accounted for here. We characterized and compared the limiting SINRs of the three schemes, in particular, using the notion of *effective interference*, which can also be used to establish if a given SINR target is feasible. The validity of the obtained results was verified via Monte Carlo simulations.

Note that we have assumed the users' channels are selected randomly. It is important to emphasize that our conclusions do *not* hold if users have been selected based on their channels, as the conditional distributions change drastically. Indeed, it is well known that for a large enough pool of users, with careful scheduling, the performance of ZF can be almost optimal.

## APPENDIX A USEFUL THEOREMS

We start by reproducing a few lemmas, which play an important role in our derivations.

**Lemma 1** (Lemma 6.3.3 in [52]). *Let  $\rho > 0$ ,  $\mathbf{A}$  and  $\mathbf{B}$   $N \times N$  matrices with  $\mathbf{B}$  Hermitian,  $\tau \in \mathbb{R}$ , and  $\mathbf{q} \in \mathbb{C}^N$ . Then*

$$\left| \text{tr} \left( \left( (\mathbf{B} + \rho \mathbf{I})^{-1} - (\mathbf{B} + \mathbf{q} \mathbf{q}^H + \rho \mathbf{I})^{-1} \right) \mathbf{A} \right) \right| \leq \frac{\|\mathbf{A}\|}{\rho} \quad (64)$$

We will also be making use of results from [53] (some themselves reproduced from [54]). [53] derives a central limit theorem for the SINR at the receiver in a multiple-access

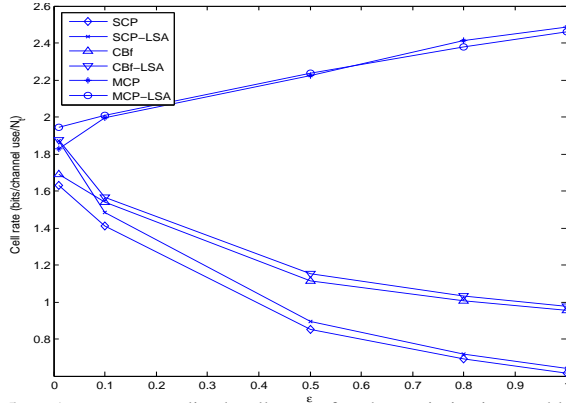


Fig. 5. Average normalized cell rates for the optimization problems for  $K=3, N=4$  and LSA rates for  $\beta=.75$ , both at SNR = 10 dB.

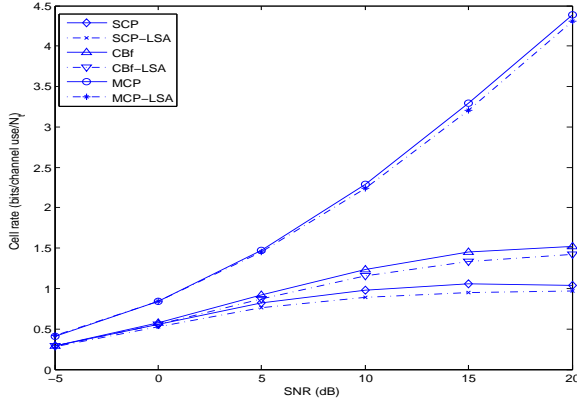


Fig. 6. Average normalized cell rates from using the asymptotically optimal beamformers to the finite system case for  $K=3, N=4$  and LSA results for  $\beta=.75$ , both with  $\epsilon=0.5$ .

MIMO system, where the dimensions of the system (number of users and the size of the received random vector) grows large: the considered asymptotic regime satisfies

$$\tilde{K} \rightarrow \infty, \liminf \frac{\tilde{K}}{\tilde{N}} > 0, \quad \limsup \frac{\tilde{K}}{\tilde{N}} < \infty, \quad (65)$$

and the quantity studied, i.e. the UL SINR, is equal to

$$\Gamma_{\tilde{K}} = \mathbf{y}^H (\mathbf{Y}\mathbf{Y}^H + \rho \mathbf{I}_{\tilde{N}})^{-1} \mathbf{y}, \quad (66)$$

where the sequence of matrices  $\Sigma(\tilde{K}) = [\mathbf{y}(\tilde{K})\mathbf{y}(\tilde{K})^H]$  is given by

$$\Sigma(\tilde{K}) = \left( \Sigma_{nk}(\tilde{K}) \right)_{n=1, k=0}^{\tilde{N}, \tilde{K}} = \left( \frac{\tilde{\sigma}_{nk}(\tilde{K})}{\sqrt{\tilde{K}}} w_{nk} \right)_{n=1, k=0}^{\tilde{N}, \tilde{K}}, \quad (67)$$

and is such that the following assumptions hold:

**A1:** the complex r.v.'s  $(w_{nk} : n \geq 1, k \geq 0)$  are i.i.d.,  $\mathbb{E}w_{10} = 0$ ,  $\mathbb{E}w_{10}^2 = 0$ ,  $\mathbb{E}|w_{10}|^2 = 1$  and  $\mathbb{E}|w_{10}|^8 < \infty$ <sup>8</sup>

**A2:** the variance profile is such that there exists a real number  $\tilde{\sigma}_{\max} < \infty$  such that

$$\sup_{\tilde{K} \geq 1} \max_{1 \leq n \leq \tilde{N}, 0 \leq k \leq \tilde{K}} |\tilde{\sigma}_{nk}(\tilde{K})| \leq \tilde{\sigma}_{\max}, \quad (68)$$

<sup>8</sup>For the channel model considered here, the entries are i.i.d.  $\mathcal{CN}(0, 1)$ , so they clearly satisfy this condition.

The dual UL SINR expressions in the present paper for  $\lambda_{kj}$ 's in each cell held at some constant value, will be of the same form as (66), and the matrices considered will satisfy assumptions **A1** and **A2**. In fact, they constitute a special case of the above model, since we consider the asymptotic regime where  $\frac{K}{N} = c$  as  $K, N \rightarrow \infty$ , and the variance profiles of the random matrices considered (see (67)) are either scaled all ones matrices, or obtained by the regular sampling of a piece-wise continuous function. Note that the expressions for asymptotic SINR on the dual UL's may equivalently be obtained from earlier results in the literature, [11] for example.

We thus reproduce below the most relevant results from [53].

**Theorem 5** (Parts 1 and 3 of Theorem 1 in [53]). *The following statements hold true.*

- Let  $(\tilde{\sigma}_{nk}^2(\tilde{K}) : 1 \leq n \leq \tilde{N}, 1 \leq k \leq \tilde{K})$  be a sequence of arrays of real numbers and consider matrices  $\mathbf{D}_k(\tilde{K})$

$$\mathbf{D}_k(\tilde{K}) = \text{diag} \left( \tilde{\sigma}_{1k}^2(\tilde{K}), \dots, \tilde{\sigma}_{\tilde{N}k}^2(\tilde{K}) \right), 0 \leq k \leq \tilde{K}. \quad (69)$$

The system of  $\tilde{N}$  functional equations

$$t_{n,\tilde{K}}(z) = \frac{1}{-z + \frac{1}{\tilde{K}} \sum_{k=1}^{\tilde{K}} \frac{\tilde{\sigma}_{nk}^2(\tilde{K})}{1 + \frac{1}{\tilde{K}} \text{tr} \mathbf{D}_k(\tilde{K}) \mathbf{T}_{\tilde{K}}(z)}}, \quad (70)$$

for  $1 \leq n \leq \tilde{N}$  and where

$$\mathbf{T}_{\tilde{K}}(z) = \text{diag} \left( t_{1,\tilde{K}}(z), \dots, t_{\tilde{N},\tilde{K}}(z) \right) \quad (71)$$

admits a unique solution  $\mathbf{T}$  among the diagonal matrices for which the  $t_{n,\tilde{K}}$  belong to class  $\mathcal{S}^9$ . Moreover, the functions admit analytical continuations over  $\mathbb{C} - [0, \infty)$  which are real and positive for  $z \in (-\infty, 0)$ .

- Assume **A1** and **A2** hold true. Consider the sequence of random matrices  $\mathbf{Y}(\tilde{K})\mathbf{Y}(\tilde{K})^H$ , where  $\mathbf{Y}_{nk} = \frac{\tilde{\sigma}_{nk}}{\sqrt{\tilde{K}}} w_{nk}$ .

For every sequence  $\mathbf{S}_K$  of  $\tilde{N} \times \tilde{N}$  diagonal matrices with

$$\sup_{\tilde{K}} \|\mathbf{S}_{\tilde{K}}\| < \infty \quad (72)$$

the following limits hold true almost surely (a.s.):

$$\lim_{\tilde{K} \rightarrow \infty} \frac{1}{\tilde{K}} \text{tr} \mathbf{S}_{\tilde{K}} (\mathbf{Q}_{\tilde{K}}(z) - \mathbf{T}_{\tilde{K}}(z)) = 0, \forall z \in \mathbb{C} - \mathbb{R}_+, \quad (73)$$

where  $\mathbf{Q}_{\tilde{K}}(z)$  denotes the resolvent of  $\mathbf{Y}(\tilde{K})\mathbf{Y}(\tilde{K})^H$ , i.e. the  $\tilde{N} \times \tilde{N}$  matrix defined by

$$\mathbf{Q}_{\tilde{K}}(z) = \left( \mathbf{Y}(\tilde{K})\mathbf{Y}(\tilde{K})^H - z \mathbf{I}_{\tilde{N}} \right)^{-1}. \quad (74)$$

**Corollary 4.** Assume **A1** and **A2** hold true. Let  $\Xi_{\tilde{K}} = \frac{1}{\tilde{K}} \text{tr} \left( \mathbf{S}_{\tilde{K}} \mathbf{Q}_{\tilde{K}}^2(-\rho) \right)$  where  $\rho \in \mathbb{R}_+$ ,  $\mathbf{S}_{\tilde{K}}$ ,  $\bar{\Xi}_{\tilde{K}} = -\frac{1}{\tilde{K}} \text{tr} \mathbf{S}_{\tilde{K}} \frac{d}{d\rho} (\mathbf{T}_{\tilde{K}}(-\rho))$ , and  $\mathbf{T}_{\tilde{K}}$  be as given by Theorem 5. Then

$$\Xi_{\tilde{K}} - \bar{\Xi}_{\tilde{K}} \xrightarrow{\tilde{K} \rightarrow \infty} 0 \text{ a.s.} \quad (75)$$

<sup>9</sup>A complex function  $t(z)$  belongs to class  $\mathcal{S}$  if  $t(z)$  is analytical in the upper half plane  $\mathbb{C}_+ = \{\text{im}(z) > 0\}$ , if  $t(z) \in \mathbb{C}_+$  for all  $z \in \mathbb{C}_+$ , and if  $\text{im} z |t(z)|$  is bounded over the upper half plane  $\mathbb{C}_+$ .

*Proof:* Differentiating (73) with respect to  $z$ , we get

$$\lim_{\tilde{K} \rightarrow \infty} \frac{1}{\tilde{K}} \text{tr}(\mathbf{S}_{\tilde{K}} \mathbf{Q}_{\tilde{K}}^2(z)) - \frac{1}{\tilde{K}} \text{tr} \mathbf{S}_{\tilde{K}} \frac{d}{dz} (\mathbf{T}_{\tilde{K}}(z)) = 0, \quad \forall z \in \mathbb{C} - \mathbb{R}_+ \quad (76)$$

$\frac{d}{dz} (\mathbf{T}_{\tilde{K}}(z))$  exists since  $t_{n,\tilde{K}}(z)$  admit analytical continuations over the range considered. This yields the result. ■

**Theorem 6** (Theorem 2 in [53]). *Let  $\bar{\Gamma}_{\tilde{K}} = \frac{1}{\tilde{K}} \text{tr}(\mathbf{D}_0(\tilde{K}) \mathbf{T}_{\tilde{K}}(-\rho))$  where  $\rho \in \mathbb{R}_+$ , and  $\mathbf{T}_{\tilde{K}}$  is given by Theorem 5. Assume A1 and A2 then*

$$\Gamma_{\tilde{K}} - \bar{\Gamma}_{\tilde{K}} \xrightarrow{\tilde{K} \rightarrow \infty} 0 \text{ a.s.} \quad (77)$$

**Proposition 4** ([53]). *Introduce the r.v.'s  $U_l = \frac{1}{\tilde{K}} \text{tr} \mathbf{D}_0 \mathbf{Q} \mathbf{D}_l \mathbf{Q}$ , for  $0 \leq l \leq \tilde{K}$ . The  $U_l$  satisfy the following system of equations:*

$$U_l = \sum_{k=1}^K c_{lk} U_k + \frac{1}{\tilde{K}} \text{tr} \mathbf{D}_0 \mathbf{D}_l \mathbf{T}^2 + \epsilon_l, 0 \leq l \leq \tilde{K}, \quad (78)$$

where

$$c_{lk} = \frac{1}{\tilde{K}} \frac{\frac{1}{\tilde{K}} \text{tr} \mathbf{D}_l \mathbf{D}_k \mathbf{T}(-\rho)^2}{\left(1 + \frac{1}{\tilde{K}} \text{tr} \mathbf{D}_k \mathbf{T}(-\rho)\right)^2}, \quad (79)$$

and the perturbations  $\epsilon_l$  satisfy  $\mathbb{E}|\epsilon_l| \leq C \tilde{K}^{-1/2}$ , where  $C$  is independent of  $l$ .

## APPENDIX B

### USEFUL ASYMPTOTIC EXPRESSIONS

In this appendix, we derive asymptotic expressions for quantities of interest in the large asymptotic analysis of SCP, CBf and MCP for the considered two cell symmetric channel model. This relies on the application of the theorems in Appendix A, noting that for our model the  $\mathbf{T}$  diagonal matrix in Theorem 5 collapses to one or two variables, as many of its entries will be equal.

#### A. SCP

Clearly, our Rayleigh channel model satisfies the conditions in Theorem 6, so that for  $\lambda_{k,j}$ 's fixed for all users in cell  $j$  at bounded<sup>10</sup>  $\lambda_j$ , the quantity of interest for any user is

$$\frac{\lambda_j}{N} \mathbf{h}_{k,j,j} \left( \frac{\lambda_j}{N} \sum_{k' \neq k} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} + \mathbf{I} \right)^{-1} \mathbf{h}_{k,j,j}^H. \quad (80)$$

From Theorem 6, this quantity will converge a.s. as  $N, K \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta$ , to<sup>11</sup>

$$\gamma^{SCP,UL}(\lambda_j) = \lambda_j t_{SCP}(-1, \lambda_j), \quad (81)$$

<sup>10</sup>This is required for the results in Appendix A to be applicable. On the other hand, if this was not the case, in the studied optimization problem, the dual objective would be unbounded and the primal unfeasible.

<sup>11</sup>The  $\mathbf{T}_{\tilde{K}}$  matrix in the theorem is simply a scaled identity with scaling factor converging to  $t_{SCP}(-1, \lambda_j)$ , as given below.

where

$$t_{SCP}(z, \lambda_j) = \frac{1}{-z + \frac{\beta \lambda_j}{1 + \lambda_j t_{SCP}(z, \lambda_j)}}. \quad (82)$$

Thus,  $\gamma^{SCP,UL}(\lambda_j)$  satisfies

$$\gamma^{SCP,UL}(\lambda_j) = \frac{\lambda_j}{1 + \frac{\beta \lambda_j}{1 + \gamma^{SCP,UL}(\lambda_j)}}. \quad (83)$$

Applying (73), we can show that

$$\frac{1}{N} \text{tr} \left[ \left( \frac{\lambda_j}{N} \sum_{k'} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} + \mathbf{I} \right)^{-1} \right] \xrightarrow{\text{a.s.}} t_{SCP}(-\rho, \lambda_j). \quad (84)$$

Furthermore, applying Corollary 4, we get that

$$\frac{1}{N} \text{tr} \left[ \left( \frac{\lambda_j}{N} \sum_{k'} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} + \mathbf{I} \right)^{-2} \right] \quad (85)$$

converges a.s. as  $N, K \rightarrow \infty$ ,  $\frac{K}{N} \rightarrow \beta$ , to

$$-\frac{d}{d\rho} t_{SCP}(-\rho, \lambda_j) \Big|_{\rho=1}. \quad (86)$$

From (82),  $t_{SCP}(-\rho, \lambda_j) = \frac{1}{\rho + \frac{\beta \lambda_j}{1 + \lambda_j t_{SCP}(-\rho, \lambda_j)}}$ , so that

$$\frac{d}{d\rho} t_{SCP}(-\rho, \lambda_j) = - \frac{t_{SCP}(-\rho, \lambda_j)}{\left[ \rho + \frac{\beta \lambda_j}{(1 + \lambda_j t_{SCP}(-\rho, \lambda_j))^2} \right]}. \quad (87)$$

#### B. CBf

Here too, we may apply Theorem 6 when  $\lambda_{k,j}$ 's for all users in cell  $j$  fixed at  $\lambda_j$ , and the values of  $\mu_j$  also held constant in both cells. For any user in cell  $j$ , we will need to characterize

$$\frac{\lambda_j}{N} \mathbf{h}_{k,j,j} (\mathbf{Y}_{k,j} \mathbf{Y}_{u,j}^H + \mu_j \mathbf{I})^{-1} \mathbf{h}_{k,j,j}^H, \quad (88)$$

where

$$\begin{aligned} & \mathbf{Y}_{k,j} \mathbf{Y}_{k,j}^H \\ &= \frac{\lambda_j}{N} \sum_{k' \neq k} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j} + \frac{\lambda_j}{N} \sum_{k'} \mathbf{h}_{k',j,j}^H \mathbf{h}_{k',j,j}. \end{aligned} \quad (89)$$

From Theorem 6, this quantity will converge a.s. as  $N, K \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta$ , to<sup>12</sup>

$$\gamma^{CBf,UL}(\mu_j, \lambda_j, \lambda_j) = \lambda_j t_{CBf}(-\mu_j, \lambda_j, \lambda_j), \quad (90)$$

where

$$\begin{aligned} & t_{CBf}(z, \lambda_j, \lambda_j) \\ &= \frac{1}{-z + \frac{\beta \lambda_j}{(1 + \lambda_j t_{CBf}(z, \lambda_j, \lambda_j))} + \frac{\beta \epsilon \lambda_j}{(1 + \epsilon \lambda_j t_{CBf}(z, \lambda_j, \lambda_j))}}. \end{aligned} \quad (91)$$

<sup>12</sup>This uses the fact that  $\frac{K-1}{K} \rightarrow 1$  and  $\frac{K}{N} \rightarrow \beta$  as  $K \rightarrow \infty$ .

Thus,  $\gamma^{CBf,UL}(\mu_j, \lambda_j, \lambda_{\bar{j}})$  satisfies

$$\gamma^{CBf,UL}(\mu_j, \lambda_j, \lambda_{\bar{j}}) = \frac{\lambda_j}{\mu_j + \frac{\beta\lambda_j}{1+\gamma^{CBf,UL}(\mu_j, \lambda_j, \lambda_{\bar{j}})} + \frac{\beta\epsilon\lambda_{\bar{j}}}{1+\epsilon\frac{\lambda_{\bar{j}}}{\lambda_j}\gamma^{CBf,UL}(\mu_j, \lambda_j, \lambda_{\bar{j}})}}. \quad (92)$$

Applying (73), we can show that

$$\frac{1}{N} \text{tr} \left( \sum_{j'=1}^2 \frac{\lambda_{j'}}{N} \sum_{k'} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} + \mu_j \mathbf{I} \right)^{-1} \xrightarrow{\text{a.s.}} t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}}). \quad (93)$$

Moreover, by Corollary 4, we can also show that

$$\frac{1}{N} \text{tr} \left( \sum_{j'=1}^2 \frac{\lambda_{j'}}{N} \sum_{k'} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} + \mu_j \mathbf{I} \right)^{-2} \quad (94)$$

converges a.s. as  $N, K \rightarrow \infty, \frac{K}{N} \rightarrow \beta$ , to

$$-\frac{d}{d\rho} t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}}) \Big|_{\rho=\mu_j}. \quad (95)$$

From (91),

$$\begin{aligned} & \frac{d}{d\rho} t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}}) \\ &= \frac{-t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}})}{\rho + \frac{\beta\lambda_j}{(1+\lambda_j t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}}))^2} + \frac{\beta\epsilon\lambda_{\bar{j}}}{(1+\epsilon\lambda_{\bar{j}} t_{CBf}(-\rho, \lambda_j, \lambda_{\bar{j}}))^2}}. \end{aligned} \quad (96)$$

### C. MCP

For  $\lambda_{k,j}$ 's in cell  $j$  equal to a constant  $\lambda_j$  and the  $\mu_j$ 's also fixed held fixed, for any user in cell  $j$ , we will need an asymptotic expression for (cf. Eq. (34))

$$\frac{\lambda_j}{N} \check{\mathbf{h}}_{k,j} \left[ \frac{\lambda_j}{N} \sum_{k' \neq k} \check{\mathbf{h}}_{k',j}^H \check{\mathbf{h}}_{k',j} + \frac{\lambda_{\bar{j}}}{N} \sum_{k'} \check{\mathbf{h}}_{k',\bar{j}}^H \check{\mathbf{h}}_{k',\bar{j}} + \mathbf{I} \right]^{-1} \check{\mathbf{h}}_{k,j}^H. \quad (97)$$

Introducing the short-hand  $\boldsymbol{\eta} = (\mu_j, \mu_{\bar{j}}, \lambda_j, \lambda_{\bar{j}})$ , applying Theorem 6, we can show that (97) converges a.s. as  $N, K \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta$ , to  $\gamma_1^{MCP,UL}(\mu_j, \mu_{\bar{j}}, \lambda_j)$  equal to

$$\gamma_1^{MCP,UL}(\boldsymbol{\eta}) = \lambda_j \left( \frac{t_{1,MCP}(-1, \boldsymbol{\eta})}{\mu_j} + \frac{\epsilon t_{2,MCP}(-1, \boldsymbol{\eta})}{\mu_{\bar{j}}} \right), \quad (98)$$

where  $t_{1,MCP}(z, \boldsymbol{\eta})$  and  $t_{2,MCP}(z, \boldsymbol{\eta})$  are given by (99).

For a user in cell  $\bar{j}$ , one can verify that (97) (replace  $j$  by  $\bar{j}$  and vice versa) converges to

$$\gamma_2^{MCP,UL}(\boldsymbol{\eta}) = \lambda_{\bar{j}} \left( \epsilon \frac{t_{1,MCP}(-1, \boldsymbol{\eta})}{\mu_j} + \frac{t_{2,MCP}(-1, \boldsymbol{\eta})}{\mu_{\bar{j}}} \right). \quad (100)$$

Plugging in (98) and (100) in the expressions for  $t_{1,MCP}(-1, \boldsymbol{\eta})$  and  $t_{2,MCP}(-1, \boldsymbol{\eta})$ , (cf. (99)), we get

$$\begin{aligned} t_{1,MCP}(-1, \boldsymbol{\eta}) &= \frac{1}{1 + \frac{\beta\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\epsilon\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}}, \\ t_{2,MCP}(-1, \boldsymbol{\eta}) &= \frac{1}{1 + \frac{\beta\epsilon\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}}. \end{aligned} \quad (101)$$

Now using (101) in (98) and (100),  $\gamma_1^{MCP,UL}(\boldsymbol{\eta})$  and  $\gamma_2^{MCP,UL}(\boldsymbol{\eta})$  will satisfy

$$\begin{aligned} & \gamma_1^{MCP,UL}(\boldsymbol{\eta}) \\ &= \lambda_j \left[ \frac{1}{\mu_j + \frac{\beta\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\epsilon\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}} + \frac{\epsilon}{\mu_{\bar{j}} + \frac{\beta\epsilon\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}} \right], \\ & \gamma_2^{MCP,UL}(\boldsymbol{\eta}) \\ &= \lambda_{\bar{j}} \left[ \frac{\epsilon}{\mu_j + \frac{\beta\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\epsilon\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}} + \frac{1}{\mu_{\bar{j}} + \frac{\beta\epsilon\lambda_j}{1+\gamma_1^{MCP,UL}(\boldsymbol{\eta})} + \frac{\beta\lambda_{\bar{j}}}{1+\gamma_2^{MCP,UL}(\boldsymbol{\eta})}} \right]. \end{aligned} \quad (102)$$

Now define

$$\mathbf{D}_{k,1} = \begin{bmatrix} \frac{\lambda_1}{\mu_1} \mathbf{I}_N & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \epsilon \frac{\lambda_1}{\mu_2} \end{bmatrix} \quad (103)$$

$$\mathbf{D}_{k,2} = \begin{bmatrix} \epsilon \frac{\lambda_2}{\mu_1} \mathbf{I}_N & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \frac{\lambda_2}{\mu_2} \end{bmatrix}. \quad (104)$$

Applying (73), we can show that

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \left( \sum_{j'=1}^2 \frac{\lambda_{j'}}{N} \sum_{k'} \check{\mathbf{h}}_{k',j'}^H \check{\mathbf{h}}_{k',j'} + \mathbf{I} \right)^{-1} \xrightarrow{\text{a.s.}} \gamma_j^{MCP,UL}(\boldsymbol{\eta}). \quad (105)$$

We also need to characterize

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \left[ \sum_{j'=1}^2 \frac{\lambda_{j'}}{N} \sum_{k'} \check{\mathbf{h}}_{k',j'}^H \check{\mathbf{h}}_{k',j'} + \mathbf{I} \right]^{-2}, \quad (106)$$

for the special case where  $\mu_j = \mu_{\bar{j}} = \mu$  and  $\lambda_j = \lambda_{\bar{j}} = \lambda$ , i.e.  $\boldsymbol{\eta} = \boldsymbol{\eta}_{sym} = [\mu, \mu, \lambda, \lambda]$ . In this case,

$$\begin{aligned} t_{1,MCP}(-\rho, \boldsymbol{\eta}_{sym}) &= t_{2,MCP}(-\rho, \boldsymbol{\eta}_{sym}) = t_{MCP}(-\rho, \boldsymbol{\eta}_{sym}) \\ &= \frac{1}{\rho + \frac{(1+\epsilon)\beta\lambda}{\mu + (1+\epsilon)\lambda t_{MCP}(-\rho, \boldsymbol{\eta}_{sym})}}. \end{aligned} \quad (107)$$

Thus,

$$\frac{d}{d\rho} t_{MCP}(-\rho, \boldsymbol{\eta}_{sym}) = - \frac{t_{MCP}(-\rho, \boldsymbol{\eta}_{sym})}{\rho + \frac{(1+\epsilon)\beta\lambda\mu}{(\mu + (1+\epsilon)\lambda t_{MCP}(-\rho, \boldsymbol{\eta}_{sym}))^2}}, \quad (108)$$

$$\begin{aligned}
t_{1,MCP}(z, \boldsymbol{\eta}) &= \frac{1}{-z + \frac{\beta \frac{\lambda_j}{\mu_j}}{1 + \frac{\lambda_j}{\mu_j} t_{1,MCP}(z, \boldsymbol{\eta}) + \frac{\epsilon}{\mu_j} \lambda_j t_{2,MCP}(z, \boldsymbol{\eta})} + \frac{\epsilon \beta \frac{\lambda_j}{\mu_j}}{1 + \frac{\lambda_j}{\mu_j} t_{1,MCP}(z, \boldsymbol{\eta}) + \frac{1}{\mu_j} \lambda_j t_{2,MCP}(z, \boldsymbol{\eta})}}, \\
t_{2,MCP}(z, \boldsymbol{\eta}) &= \frac{1}{-z + \frac{\beta \epsilon \frac{\lambda_j}{\mu_j}}{1 + \frac{1}{\mu_j} \lambda_j t_{1,MCP}(z, \boldsymbol{\eta}) + \frac{\lambda_j}{\mu_j} t_{2,MCP}(z, \boldsymbol{\eta})} + \frac{\beta \frac{\lambda_j}{\mu_j}}{1 + \frac{\epsilon}{\mu_j} \lambda_j t_{1,MCP}(z, \boldsymbol{\eta}) + \frac{\lambda_j}{\mu_j} t_{2,MCP}(z, \boldsymbol{\eta})}}. \tag{99}
\end{aligned}$$

and, by Corollary 4, (106) converges a.s. as  $N, K \rightarrow \infty$  with  $\frac{K}{N} \rightarrow \beta$  to

$$-\frac{(1+\epsilon)\lambda}{\mu} \frac{d}{d\rho} t_{MCP}(-\rho, \boldsymbol{\eta}_{sym}) \Big|_{\rho=1}. \tag{109}$$

Finally, we will need to characterize

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \mathbf{D}_{k',j'} \mathbf{A} \tag{110}$$

with  $\mathbf{A}$  as defined in (194), where in  $\lambda_1$  and  $\lambda_2$  in  $\mathbf{D}_{k,1}$ ,  $\mathbf{D}_{k',2}$  are equal to  $\bar{\lambda}$  for all  $k \leq K$ . For any sequence of diagonal matrices  $\mathbf{S}$  with bounded diagonal entries, and taking into account the fact that the  $\mathbf{D}_{k,j}$ 's are equal for all users in the same cell, define  $V_j(\mathbf{S}) = \frac{1}{N} \text{tr} \mathbf{S} \mathbf{A} \mathbf{D}_{1,j} \mathbf{A}$ , for  $j = 1, 2$ . Applying Proposition 4, we get that

$$\begin{aligned}
V_1(\mathbf{S}) &= \beta \frac{t_{MCP}^2 \bar{\lambda}^2 (1 + \epsilon^2)}{(1 + (1 + \epsilon) \bar{\lambda} t_{MCP})^2} V_1(\mathbf{S}) \\
&+ \beta \frac{2\epsilon t_{MCP}^2 \bar{\lambda}^2}{(1 + (1 + \epsilon) \bar{\lambda} t_{MCP})^2} V_2(\mathbf{S}) \\
&+ \frac{t_{MCP}^2}{N} \text{tr} \mathbf{S} \mathbf{D}_{1,1} + \eta_1, \tag{111}
\end{aligned}$$

$$\begin{aligned}
V_2(\mathbf{S}) &= \beta \frac{2\epsilon t_{MCP}^2 \bar{\lambda}^2}{(1 + (1 + \epsilon) \bar{\lambda} t_{MCP})^2} V_1(\mathbf{S}) \\
&+ \beta \frac{t_{MCP}^2 \bar{\lambda}^2 (1 + \epsilon^2)}{(1 + (1 + \epsilon) \bar{\lambda} t_{MCP})^2} V_2(\mathbf{S}) \\
&+ \frac{t_{MCP}^2}{N} \text{tr} \mathbf{S} \mathbf{D}_{1,2} + \eta_2, \tag{112}
\end{aligned}$$

where  $\eta_1$  and  $\eta_2$  satisfy  $\mathbb{E}|\eta_j| \leq CN^{-1/2}$ ,  $j = 1, 2$ , and  $t_{MCP} = t_{MCP}(1, 1, \bar{\lambda}, \bar{\lambda})$ , for some constant  $C$ .

For  $\mathbf{S} = \mathbf{D}_{1,1}$ , we obtain (113) and (114), where  $\phi_1$  and  $\phi_2$  denote deviation terms such that  $\mathbb{E}|\phi_j| \leq CN^{-1/2}$ .

Note that by definition,  $V_1(\mathbf{D}_{1,2}) = V_2(\mathbf{D}_{1,1})$ , and  $V_2(\mathbf{D}_{1,2}) = V_1(\mathbf{D}_{1,1})$ .

## APPENDIX C

### A SIMPLE MONOTONICITY RESULT

Let  $I(\boldsymbol{\lambda})$  be a standard interference function for the UL, in the sense of Yates [27], where we denote the transmit power vector by  $\boldsymbol{\lambda}$ . Suppose it is of the form

$$I_k(\boldsymbol{\lambda}) = \gamma_k F_k(\boldsymbol{\lambda}) \quad k = 1, 2, \dots, K \tag{115}$$

where  $\gamma_k$  is the SINR target for user  $k$ , and  $F(\boldsymbol{\lambda})$  is a standard interference function. The vector  $\boldsymbol{\lambda}$  is called *feasible* if

$$\lambda_k \geq \gamma_k F_k(\boldsymbol{\lambda}) \quad k = 1, 2, \dots, K. \tag{116}$$

It is shown in [27], Theorem 1, that if a feasible solution exists, then function  $I$  has a unique, positive, fixed point. [27] also shows that, starting at any power vector  $\boldsymbol{\lambda}$ , the iterative power control  $I^n(\boldsymbol{\lambda})$ ,  $n = 1, 2, \dots$  converges to it. Using the fixed-point powers, all users achieve exactly their SINR target.

Now consider two different vectors of SINR targets  $\boldsymbol{\gamma}^{(1)}$  and  $\boldsymbol{\gamma}^{(2)}$ , and let  $\boldsymbol{\lambda}^{(1)}$  and  $\boldsymbol{\lambda}^{(2)}$  denote the corresponding fixed points. Lemma 2 is a simple corollary of Lemma 1 in [27].

**Lemma 2.** *If  $\boldsymbol{\gamma}^{(1)} \leq \boldsymbol{\gamma}^{(2)}$  then  $\boldsymbol{\lambda}^{(1)} \leq \boldsymbol{\lambda}^{(2)}$ .*

*Proof:*  $\boldsymbol{\lambda}^{(2)}$  is feasible for the power control problem with SINR targets given by  $\boldsymbol{\gamma}^{(1)}$ . By [27], Lemma 1,  $I^n(\boldsymbol{\lambda}^{(2)})$ ,  $n = 1, 2, \dots$  is a monotone decreasing sequence of feasible power vectors that converges to  $\boldsymbol{\lambda}^{(1)}$ . ■

## APPENDIX D PROOF OF THEOREM 1

Throughout this section, let

$$\mathbf{A}_j = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{l=1}^K \mathbf{h}_{l,j}^H \mathbf{h}_{l,j} \right)^{-1} \tag{117}$$

$$\mathbf{A}_{k,j} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{l \neq k} \mathbf{h}_{l,j}^H \mathbf{h}_{l,j} \right)^{-1} \tag{118}$$

$$\mathbf{A}_{k,k',j} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{l \neq (k,k')} \mathbf{h}_{l,j}^H \mathbf{h}_{l,j} \right)^{-1}, \tag{119}$$

where  $\bar{\lambda} > 0$  will be defined later.

### Asymptotic analysis of the dual problem

We start by considering the dual problem at each of the BSs: this will yield the asymptotically optimal dual variables and beamforming directions in both dual and primal problems. Note that even though the value of the dual objective function in one cell depends on the primal beamforming decisions through the  $\sigma_{k,j}$ s in (14), the optimal dual variables themselves are fully determined by the constraints and will be the unique strictly positive solutions to (18). Since the analysis is identical in both cells, without loss of generality, assume the cell index  $j$  is  $j = 1$ . Assume also that  $\beta \left( \frac{\gamma}{1 + \gamma} + \epsilon \gamma \right) < 1$ . As noted in Section V, we cannot immediately apply standard large system analysis to (17) as optimal dual variables  $\lambda_{k,1}$ 's are not independent of the channel vectors  $(\mathbf{h}_{k,1,1})_{k=1}^K$ .

Rather than directly analyze the asymptotic performance of the SCP system using *optimal*  $\lambda_{k,1}$ 's, consider any constant

$$V_1(\mathbf{D}_{1,1}) = t_{MCP}^2 \bar{\lambda}^2 (1+\epsilon)^2 \frac{\frac{1+\epsilon^2}{(1+\epsilon)^2} - \beta \frac{(1-\epsilon)^2 \bar{\lambda}^2 t_{MCP}^2}{(1+(1+\epsilon)\bar{\lambda}t_{MCP})^2}}{\left[1 - \beta \frac{(1+\epsilon)^2 \bar{\lambda}^2 t_{MCP}^2}{(1+(1+\epsilon)\bar{\lambda}t_{MCP})^2}\right] \left[1 - \beta \frac{(1-\epsilon)^2 \bar{\lambda}^2 t_{MCP}^2}{(1+(1+\epsilon)\bar{\lambda}t_{MCP})^2}\right]} + \phi_1 \quad (113)$$

$$V_2(\mathbf{D}_{1,1}) = t_{MCP}^2 \bar{\lambda}^2 (1+\epsilon)^2 \frac{\frac{2\epsilon}{(1+\epsilon)^2}}{\left[1 - \beta \frac{(1+\epsilon)^2 \bar{\lambda}^2 t_{MCP}^2}{(1+(1+\epsilon)\bar{\lambda}t_{MCP})^2}\right] \left[1 - \beta \frac{(1-\epsilon)^2 \bar{\lambda}^2 t_{MCP}^2}{(1+(1+\epsilon)\bar{\lambda}t_{MCP})^2}\right]} + \phi_2. \quad (114)$$

$\bar{\lambda} > 0$  (later, we will assign it a particular value, but for now it is arbitrary) and consider the large system regime in which all users have the same transmit UL power of  $\bar{\lambda}/N$ , as  $N, K \rightarrow \infty$ , with fixed ratio  $K/N = \beta$ . This is a virtual UL with noise of unit power, so the SINR for user  $k$  is given by

$$\frac{\bar{\lambda}}{N} \mathbf{h}_{k,1,1} \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{k' \neq k} \mathbf{h}_{k',1,1}^H \mathbf{h}_{k',1,1} \right)^{-1} \mathbf{h}_{k,1,1}^H. \quad (120)$$

Appendix B-A shows how this converges a.s. to a constant in the considered asymptotic regime.

In particular, if  $\delta$  is small enough so that  $\beta \frac{\gamma + \delta}{1 \pm \gamma + \delta} < 1$ , and we assign the following particular value to  $\bar{\lambda}$ :

$$\bar{\lambda}(\delta) := \frac{\gamma + \delta}{1 - \beta \frac{\gamma + \delta}{1 + \gamma + \delta}}, \quad (121)$$

which we denote by  $\bar{\lambda}(\delta)$  to make explicit its dependence on the parameter  $\delta$ , then (120) will converge a.s. to  $\gamma + \delta$ . Theorem 3 in [53] shows that for this (suboptimal) system, the value of (120) is

$$\gamma_k(\delta) = \gamma + \delta + \mathcal{O}\left(\sqrt{\frac{1}{N}}\right), \quad \forall k = 1, 2, \dots, K, \quad (122)$$

where  $\gamma_k(\delta)$  denotes the dual UL SINR of user  $k$  under this suboptimal power allocation, and the last term on the right hand side is  $1/\sqrt{N}$  times a r.v. that converges weakly to a zero-mean Gaussian distributed r.v. This holds since  $\frac{K}{N} \rightarrow \beta < \infty$ , for the considered channel model and in the notation of the theorem,  $\Gamma_{\bar{K}}$  and  $\Theta_{\bar{K}}$  will converge a.s. to bounded limits.

Denote by  $\bar{\lambda}(\delta)$  the vector of UL powers in this suboptimal system, where all entries have the same value  $\bar{\lambda}(\delta)$ . For  $\delta > 0$ , it follows from (122) that for any particular  $k$ ,  $\mathbb{P}(\gamma_k(-\delta) > \gamma$  or  $\gamma_k(\delta) < \gamma)$  decays to 0 exponentially in  $N$  (it is a large deviation event). Applying the union bound, we obtain that

$$\gamma_k(-\delta) \leq \gamma \leq \gamma_k(\delta) \quad \forall k = 1, 2, \dots, K \quad (123)$$

will hold with probability tending to 1 as  $N \uparrow \infty$ . Let  $\gamma(-\delta)$  and  $\gamma(\delta)$  denote the vectors grouping the left and right-hand sides of (123), respectively.

Denote by  $\boldsymbol{\lambda}$  the vector of *optimal* UL powers for the dual problem, which achieves SINR of  $\gamma$  for each user. On the other hand,  $\bar{\boldsymbol{\lambda}}(-\delta)$  and  $\bar{\boldsymbol{\lambda}}(\delta)$  (defined above) are the vectors of *optimal* UL powers for the virtual UL problem with target SINR vectors  $\gamma(-\delta)$  and  $\gamma(\delta)$ , respectively (instead of the

all- $\gamma$  vector in the original dual problem). Thus, by Lemma 2 in Appendix C,

$$\bar{\boldsymbol{\lambda}}(-\delta) \leq \boldsymbol{\lambda} \leq \bar{\boldsymbol{\lambda}}(\delta) \quad (124)$$

will hold with probability tending to 1 as  $N \uparrow \infty$ . Since this holds for any sufficiently small  $\delta > 0$ , the empirical distribution of the components of the optimal  $\boldsymbol{\lambda}$  will converge weakly to the constant  $\bar{\lambda}$ , given in (39). Let (39) provide the particular value of  $\bar{\lambda}$  in the rest of this section. This establishes the asymptotic optimality of having the  $\lambda_{k,j}$ 's in both cells all equal to  $\bar{\lambda}$ , and consequently that of the beamforming vectors

$$\hat{\mathbf{w}}_{k,j} = \mathbf{A}_{k,j} \mathbf{h}_{k,j}^H. \quad (125)$$

By UL-DL duality and from the KKT conditions, these are, up to a scale factor, also the optimal DL beamforming vectors [29]. Thus, this analysis shows that, asymptotically, the DL beamforming *directions* in one cell do not depend on the beamforming *directions* used in the other cell, although it is clear that the optimal DL power levels (which modulate the beamforming directions) *will* depend on the power levels used in the other cell, even in the limit. Finally, also note that the dual objective function value  $\frac{\bar{\lambda}}{N} \sum_{k=1}^K \sigma_{k,j}^2$  is an upper bound on the primal objective function, i.e. on the total transmit power in the cell; denote the latter by  $\bar{P}_j$  for cell  $j$ .

#### Asymptotic analysis of the primal problems

We now turn to the DL primal problems, and fix the beamforming directions in both cells to be those given by (125). Thus, only the DL power levels  $p_{k,j}$  in (19) need to be determined. These must satisfy, for all users in both cells,

$$p_{k,j} = \frac{\sigma_{k,j}^2 + \sum_{k' \neq k} \frac{p_{k',j}}{N} \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2}}{\frac{1}{N\gamma} \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{k,j}|^2}{\|\hat{\mathbf{w}}_{k,j}\|^2}}, \quad (126)$$

$$\sigma_{k,j}^2 = \sigma^2 + \sum_{k'=1}^K \frac{p_{k',j}}{N} \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2}. \quad (127)$$

Assuming feasibility, one can verify using similar standard interference function arguments as for the dual problem, that the set of equations in both cells will have a unique power minimizing solution [26].

As was the case with asymptotic analysis of the dual problem, it is easier to first fix the DL powers to constants and study the resulting limiting regime. Thus, assume  $p_{k,j}$ 's in cell  $j$  are held fixed at a constant value  $\bar{p}_j$   $j = 1, 2$ , and note that this implies that  $\bar{P}_j = \beta \bar{p}_j$ . After analyzing this limiting regime, we optimize the choice of the constants,  $\bar{p}_1$  and  $\bar{p}_2$



and finally show that the optimal constants are asymptotically optimal with respect to the primal optimization problem (6).

Under the regime in which  $p_{kj} = \bar{p}_j$  for all  $k$  in cell  $j$ , the following lemmas hold.

**Lemma 3.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{kj} \mathbf{h}_{k,j,j}^H$ , such that  $\bar{\lambda} = \frac{\gamma}{1-\beta\frac{\gamma}{1+\gamma}}$ , the following holds, as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} = \beta$ :

$$\max_{j=1,2,k \leq K} \left| \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{kj}|^2}{N \|\hat{\mathbf{w}}_{kj}\|^2} - \left[ 1 - \frac{\beta\gamma^2}{(1+\gamma)^2} \right] \right| \xrightarrow{a.s.} 0 \quad (128)$$

$$\max_{j=1,2,k,l \leq K} \left| \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{lj}|^2}{\|\hat{\mathbf{w}}_{lj}\|^2} - \epsilon \right| \xrightarrow{a.s.} 0 \quad (129)$$

$$\max_{j=1,2,k,l \leq K, l \neq k} \left| \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{lj}|^2}{\|\hat{\mathbf{w}}_{lj}\|^2} - \frac{1}{(1+\gamma)^2} \right| \xrightarrow{a.s.} 0. \quad (130)$$

*Proof:* Applying Lemma 5.1 in [55],

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \|\hat{\mathbf{w}}_{kj}\|^2 - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{k,j}^2 \right| \xrightarrow{a.s.} 0 \quad (131)$$

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{kj} - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{k,j} \right| \xrightarrow{a.s.} 0, \quad (132)$$

where<sup>13</sup>  $\mathbf{D}_{k,j,j} = \bar{\lambda} \mathbf{I}$ . Later on in this section,  $\mathbf{D}_{k,j,\bar{j}} = \epsilon \bar{\lambda} \mathbf{I}$ . Now applying Lemma 1 twice to (131) and once in (132), we get

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \|\hat{\mathbf{w}}_{kj}\|^2 - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{k,j}^2 \right| \xrightarrow{a.s.} 0 \quad (133)$$

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{kj} - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{k,j} \right| \xrightarrow{a.s.} 0. \quad (134)$$

Now consider the interference terms, with  $i = 1, 2$ , and  $(i, l) \neq (k, j)$ ,

$$\frac{|\mathbf{h}_{k,j,i} \hat{\mathbf{w}}_{li}|^2}{\|\hat{\mathbf{w}}_{li}\|^2} = \frac{1}{\bar{\lambda}} \frac{\bar{\lambda}^2 \mathbf{h}_{k,j,i} \mathbf{A}_{l,i} \mathbf{h}_{l,i,i}^H \mathbf{h}_{l,i,i} \mathbf{A}_{l,i} \mathbf{h}_{k,j,i}^H}{\bar{\lambda} \|\hat{\mathbf{w}}_{li}\|^2}. \quad (135)$$

Two different cases arise, depending on whether  $i = j$  or  $i = \bar{j}$ . In the latter case, we may apply Lemma 5.1 in [55] to the numerator of the right-hand side of (135),

$$\max_{k,l \leq K} \left| \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{l,\bar{j},\bar{j}}^H \mathbf{h}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{k,j,\bar{j}}^H - \frac{\bar{\lambda}}{N} \mathbf{h}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{l,\bar{j},\bar{j}}^H \right| \xrightarrow{a.s.} 0. \quad (136)$$

Applying Lemma 5.1 in [55] once again yields,

$$\max_{k,l \leq K} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{l,\bar{j},\bar{j}}^H - \frac{1}{N} \text{tr} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} \right| \xrightarrow{a.s.} 0. \quad (137)$$

Finally applying Lemma 1 twice, we get

$$\begin{aligned} & \left| \frac{1}{N} \text{tr} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} - \frac{1}{N} \text{tr} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{\bar{j}} \right| \\ & \leq \frac{1}{N} \|\mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}}\| \leq \frac{\|\mathbf{D}_{l,\bar{j},\bar{j}}\| \|\mathbf{A}_{l,\bar{j}}\| \|\mathbf{D}_{k,j,\bar{j}}\|}{N} \\ & \leq \frac{\epsilon \bar{\lambda}^2}{N}, \end{aligned} \quad (138)$$

$$\begin{aligned} & \left| \frac{1}{N} \text{tr} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{\bar{j}} - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{\bar{j}} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{\bar{j}} \right| \\ & \leq \frac{1}{N} \|\mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{\bar{j}} \mathbf{D}_{k,j,\bar{j}}\| \leq \frac{\epsilon \bar{\lambda}^2}{N}. \end{aligned} \quad (139)$$

Thus,

$$\begin{aligned} & \max_{k,l \leq K} \left| \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{l,\bar{j},\bar{j}}^H \mathbf{h}_{l,\bar{j},\bar{j}} \mathbf{A}_{l,\bar{j}} \mathbf{h}_{k,j,\bar{j}}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{\bar{j}} \mathbf{D}_{l,\bar{j},\bar{j}} \mathbf{A}_{\bar{j}} \right| \xrightarrow{a.s.} 0. \end{aligned} \quad (140)$$

When  $i$  in (135) is the same as  $j$ , we cannot apply Lemma 5.1 in [55] directly, since  $\mathbf{A}_{l,j}$  and  $\mathbf{h}_{k,j,j}$  are not independent. Thus, we apply the matrix inversion lemma first to get

$$\begin{aligned} & \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,j} \mathbf{A}_{l,j} \mathbf{h}_{l,j,j}^H \mathbf{h}_{l,j,j} \mathbf{A}_{l,j} \mathbf{h}_{k,j,j}^H \\ & = \frac{\bar{\lambda}^2 \mathbf{h}_{k,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{l,j,j}^H \mathbf{h}_{l,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{k,j,j}^H}{\left( 1 + \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{k,j,j}^H \right)^2}. \end{aligned} \quad (141)$$

We can now consider the numerator in (141) and show that

$$\begin{aligned} & \max_{k,l \leq K, (k \neq l)} \left| \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{l,j,j}^H \mathbf{h}_{l,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{k,j,j}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{\bar{j}} \mathbf{D}_{l,j,j} \mathbf{A}_{\bar{j}} \right| \xrightarrow{a.s.} 0. \end{aligned} \quad (142)$$

We can also show that

$$\max_{k,l \leq K, (k \neq l)} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,j} \mathbf{A}_{k,l,j} \mathbf{h}_{k,j,j}^H - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{\bar{j}} \right| \xrightarrow{a.s.} 0. \quad (143)$$

The proof of the lemma is concluded by using the limits of the trace terms as derived in Appendix B-A, with  $\lambda_j = \bar{\lambda}$ , noting that for this specific case, all the  $\mathbf{D}_{k,j,j}$  and  $\mathbf{D}_{k,j,\bar{j}}$  matrices are equal, and are simply scaled identities, so that

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{\bar{j}} = \frac{\bar{\lambda}}{N} \text{tr} \mathbf{A}_{\bar{j}} \xrightarrow{a.s.} \gamma \quad (144)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{\bar{j}}^2 = \frac{\bar{\lambda}}{N} \text{tr} \mathbf{A}_{\bar{j}}^2 \xrightarrow{a.s.} \frac{\gamma}{1 + \frac{\beta \bar{\lambda}}{(1+\gamma)^2}} = \frac{1}{\bar{\lambda}} \frac{\gamma^2}{1 - \frac{\beta \gamma^2}{(1+\gamma)^2}} \quad (145)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_{\bar{j}} \mathbf{D}_{l,j,j} \mathbf{A}_{\bar{j}} = \frac{\bar{\lambda}^2}{N} \text{tr} \mathbf{A}_{\bar{j}}^2 \xrightarrow{a.s.} \frac{\gamma^2}{1 - \frac{\beta \gamma^2}{(1+\gamma)^2}} \quad (146)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j,\bar{j}} \mathbf{A}_{\bar{j}} \mathbf{D}_{l,j,\bar{j}} \mathbf{A}_{\bar{j}} = \frac{\epsilon \bar{\lambda}^2}{N} \text{tr} \mathbf{A}_{\bar{j}}^2 \xrightarrow{a.s.} \epsilon \frac{\gamma^2}{1 - \frac{\beta \gamma^2}{(1+\gamma)^2}}. \quad (147)$$

<sup>13</sup>We introduce the  $\mathbf{D}_{k,j,i}$  matrices to allow for more general formulations. ■

**Lemma 4.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{kj} \mathbf{h}_{k,j,j}^H$ , such that  $\bar{\lambda} = \frac{\gamma}{1-\beta\frac{\gamma}{1+\gamma}}$ , and with  $p_{kj} = \bar{p}_j$  for  $k = 1, \dots, K$ ,  $j = 1, 2$ , it follows that, with probability 1,

$$\sigma_{k,j}^2 \xrightarrow{a.s.} \sigma^2 + \epsilon \bar{P}_j, \quad (148)$$

$\bar{P}_j$  denotes the total transmit power of BS  $j$ , i.e.  $\bar{P}_j = \beta \bar{p}_j$ .

*Proof:*

Since DL  $p_{kj}$ s are all fixed to the same value,  $\bar{p}_j$ , we have

$$\sigma_{k,j}^2 = \sigma^2 + \frac{\bar{p}_j}{N} \sum_{k=1}^K \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k,j}|^2}{\|\hat{\mathbf{w}}_{k,j}\|^2}. \quad (149)$$

Applying (129) of Lemma 3, for  $k = 1, \dots, K$ ,  $j = 1, 2$ ,

$$\sigma_{k,j}^2 = \sigma^2 + \epsilon \bar{P}_j + o(1). \quad (150)$$

We conclude from Lemma 4 that  $\sigma_{k,j}^2$  is asymptotically independent of the user index  $k$ , in the regime considered ( $p_{kj} = \bar{p}_j$ ) and therefore write the limiting value as  $\sigma_j^2$ . This deals with the asymptotics of the RHS of (127), and we now provide two lemmas to deal with those of the RHS of (126).

**Lemma 5.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \mathbf{h}_{k,j,j}^H$ , such that  $\bar{\lambda} = \frac{\gamma}{1-\beta\frac{\gamma}{1+\gamma}}$ , and with  $p_{kj} = \bar{p}_j$  for  $k = 1, \dots, K$ ,  $j = 1, 2$ . The following holds for any  $k'$  in cell  $j$ :

$$\sum_{k' \neq k} \frac{p_{k',j}}{N} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2} \xrightarrow{a.s.} \bar{p}_j \frac{\beta}{(1+\gamma)^2} = \frac{\bar{P}_j}{(1+\gamma)^2}, \quad (151)$$

as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} = \beta$ ;  $\bar{P}_j = \beta \bar{p}_j$ .

*Proof:*

$$\begin{aligned} \sum_{k' \neq k} \frac{p_{k',j}}{N} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2} &= \frac{\bar{p}_j}{N} \sum_{k' \neq k} \frac{|\mathbf{h}_{k,j,j} \hat{\mathbf{w}}_{k',j}|^2}{\|\hat{\mathbf{w}}_{k',j}\|^2} \\ &= \frac{K-1}{N} \frac{\bar{p}_j}{(1+\gamma)^2} + o(1), \end{aligned} \quad (152)$$

where we used (130) in Lemma 3. Noting that as  $K, N \rightarrow \infty$ ,  $\frac{K-1}{N} \rightarrow \beta$  completes the proof. ■

To find the *minimal* pair of constants  $(\bar{p}_1, \bar{p}_2)$  for the two cells, we therefore solve the following set of equations,

$$\bar{p}_1 = \frac{\gamma}{1 - \frac{\beta\gamma^2}{(1+\gamma)^2}} \left[ \sigma^2 + \epsilon \beta \bar{p}_2 + \bar{p}_1 \frac{\beta}{(1+\gamma)^2} \right], \quad (153)$$

$$\bar{p}_2 = \frac{\gamma}{1 - \frac{\beta\gamma^2}{(1+\gamma)^2}} \left[ \sigma^2 + \epsilon \beta \bar{p}_1 + \bar{p}_2 \frac{\beta}{(1+\gamma)^2} \right], \quad (154)$$

to obtain

$$\bar{p}_1 = \bar{p}_2 = \bar{p} = \frac{\sigma^2 \gamma}{1 - \frac{\beta\gamma}{(1+\gamma)} - \epsilon \beta \gamma}. \quad (155)$$

This implies that

$$\bar{P}_1 = \bar{P}_2 = \bar{P} := \beta \bar{p}. \quad (156)$$

Note that such a choice of transmit powers is guaranteed to meet the SINR constraints as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} = \beta$ .

We confirm the asymptotic optimality of this deterministic power allocation, together with the beamforming directions

found from analysis of the dual problem, by verifying that the duality gap is tending to zero in both cells. Indeed, in cell  $j$ , the primal objective value converges to  $\bar{P}$ , whereas the dual objective value converges to  $\beta \bar{\lambda} (\sigma^2 + \epsilon \bar{P})$ . Recalling the definition of  $\bar{\lambda}$ , one can easily verify that these two quantities are the same, thereby completing the proof.

Equation (155) shows that the coupled primal problems have a solution when effective bandwidth condition  $\beta \left( \frac{\gamma}{1+\gamma} + \epsilon \gamma \right) < 1$  is satisfied. Conversely, if this condition does not hold, there is asymptotically no feasible solution to the coupled primal problems, in the limit as  $N \uparrow \infty$ . The latter observation follows from monotonicity: The optimal powers are increasing functions of  $\gamma$ , but as the denominator of (155) decreases to zero, the optimal total power from either BS,  $\bar{P}$ , must diverge to infinity, and for higher values of  $\gamma$  there can be no feasible solution.

## APPENDIX E PROOF OF THEOREM 2

Throughout this section, let

$$\mathbf{A}_j = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{i=1}^2 \sum_{l=1}^K \mathbf{h}_{l,i,j}^H \mathbf{h}_{l,i,j} \right)^{-1} \quad (157)$$

$$\mathbf{A}_{k,j} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{i=1}^2 \sum_{(l,i) \neq (j,k)} \mathbf{h}_{l,i,j}^H \mathbf{h}_{l,i,j} \right)^{-1} \quad (158)$$

$$\mathbf{A}_{k,j,k',j',j} = \left( \mathbf{I} + \frac{\bar{\lambda}}{N} \sum_{i=1}^2 \sum_{(l,i) \neq (j,k), (j',k')} \mathbf{h}_{l,i,j}^H \mathbf{h}_{l,i,j} \right)^{-1}, \quad (159)$$

where  $\bar{\lambda} > 0$  will be defined later.

Moreover, the following lemma will be useful in proving the theorem.

**Lemma 6.** Let  $(\mu_1, \mu_2)$  satisfy  $0 \leq \mu_i \leq 2$ ,  $i = 1, 2$ , and define the function  $F(\bar{\lambda}_1, \bar{\lambda}_2)$  by

$$\begin{aligned} F(\bar{\lambda}_1, \bar{\lambda}_2) &= (F_1(\bar{\lambda}_1, \bar{\lambda}_2), F_2(\bar{\lambda}_1, \bar{\lambda}_2)), \\ F_j(\bar{\lambda}_1, \bar{\lambda}_2) &= \gamma \left( \mu_j + \frac{\beta}{1+\gamma} \bar{\lambda}_j + \frac{\beta \epsilon}{1 + \frac{\bar{\lambda}_j}{\bar{\lambda}_2} \gamma \epsilon} \bar{\lambda}_j \right) \quad j = 1, 2. \end{aligned} \quad (160)$$

Then

- i)  $F$  is an interference function [27].
- ii) If  $\beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon \gamma}{1+\epsilon \gamma} \right) < 1$  then there exists a unique solution to the fixed point equation

$$(\bar{\lambda}_1, \bar{\lambda}_2) = F(\bar{\lambda}_1, \bar{\lambda}_2). \quad (161)$$

*Proof:* (i) is easily verified. For (ii), let the value  $\tilde{\lambda}$  be

$$\tilde{\lambda} = \frac{2\gamma}{1 - \beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon \gamma}{1+\epsilon \gamma} \right)}. \quad (162)$$

Since  $\mu_1, \mu_2 \leq 2$ , it is easy to verify that

$$F(\tilde{\lambda}, \tilde{\lambda}) \leq (\tilde{\lambda}, \tilde{\lambda}) \text{ componentwise} \quad (163)$$

and hence the inequality  $(\lambda_1, \lambda_2) \geq F(\lambda_1, \lambda_2)$  has a feasible solution. It follows from Theorem 1 in [27] that (160) has a unique fixed point  $(\bar{\lambda}_1, \bar{\lambda}_2)$ . The latter depends on  $(\mu_1, \mu_2)$ . ■

#### Asymptotic analysis of the dual problem

Assume that  $\beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right) < 1$ , let  $(\mu_1, \mu_2)$  be feasible for the dual (22) i.e.  $\mu_1, \mu_2 \geq 0$  and  $\mu_1 + \mu_2 = 2$ . Consider the suboptimal dual power vector that assigns power level  $\bar{\lambda}_j/N$  to all the users in cell  $j$ ,  $j = 1, 2$ , respectively. The dual UL SINR is equal to

$$\frac{\bar{\lambda}_j}{N} \mathbf{h}_{k,j,j} \left[ \mu_j \mathbf{I} + \sum_{j'=1}^2 \frac{\bar{\lambda}_{j'}}{N} \sum_{k', (j', k') \neq (j, k)} \mathbf{h}_{k',j',j}^H \mathbf{h}_{k',j',j} \right]^{-1} \mathbf{h}_{k,j,j}^H. \quad (164)$$

We show in Appendix B-B that this quantity converges a.s. to a constant equal to  $\bar{\lambda}_j t_{CBf}(-\mu_j, \bar{\lambda}_j, \bar{\lambda}_j)$  (cf. Eq. (91)).

Letting  $(\bar{\lambda}_1, \bar{\lambda}_2)$  be the unique solution to (161),  $\bar{\lambda}_j t_{CBf}(-\mu_j, \bar{\lambda}_j, \bar{\lambda}_j)$  will be equal to  $\gamma$  for  $j = 1, 2$ . Thus, with this set of suboptimal dual UL powers, all SINR's (cf. Eq. (164)) converge to  $\gamma$ . Note that the condition that  $0 \leq \mu_i \leq 2$ ,  $i = 1, 2$ , is weaker than the condition that  $(\mu_1, \mu_2)$  satisfy the dual feasibility condition that  $\mu_1 + \mu_2 = 2$ , but it certainly includes this condition.

We now arrive at a particular choice of  $(\mu_1, \mu_2)$  by solving the following optimization problem:

$$\max_{\mu_1, \mu_2 \geq 0} \quad \beta \sigma^2 (\bar{\lambda}_1 + \bar{\lambda}_2) \quad (165)$$

$$\text{s.t.} \quad \mu_1 + \mu_2 = 2 \quad (166)$$

$$(\bar{\lambda}_1, \bar{\lambda}_2) \text{ is the unique fixed point of (161).} \quad (167)$$

Adding up the two equations in (161), and taking account of (166), we obtain

$$\begin{aligned} & \left( 1 - \frac{\beta}{1+\gamma} \right) (\bar{\lambda}_1 + \bar{\lambda}_2) \\ &= 2\gamma + \beta\epsilon \left( \frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\lambda}_1 + \gamma\epsilon \bar{\lambda}_2} + \frac{\bar{\lambda}_1 \bar{\lambda}_2}{\bar{\lambda}_2 + \gamma\epsilon \bar{\lambda}_1} \right). \end{aligned} \quad (168)$$

Let  $t = \bar{\lambda}_1 + \bar{\lambda}_2$  for  $(\bar{\lambda}_1, \bar{\lambda}_2)$  optimal. For fixed  $t$ , the  $(\bar{\lambda}_1, \bar{\lambda}_2)$  that maximize the RHS of (168) can be verified to be  $(t/2, t/2)$ . The optimal  $\bar{\lambda}_1$  and  $\bar{\lambda}_2$  must thus be equal; we denote the common value by  $\bar{\lambda}$ . (161) then implies  $\mu_1 = \mu_2 = 1$ . As a result,  $\bar{\lambda}$  is as given in (45). We conclude that the optimal choice (with respect to the optimization problem (165)-(167)) is to take  $(\mu_1, \mu_2) = (1, 1)$ .

Note also that if we fix  $(\mu_1, \mu_2) = (1, 1)$  and use the corresponding deterministic  $(\bar{\lambda}_1, \bar{\lambda}_2) = (\bar{\lambda}, \bar{\lambda})$  (with  $\bar{\lambda}$  satisfying (45)) to generate dual UL powers in each cell, then the dual objective function will converge to the solution of the optimization problem (165)-(167).

Now consider the optimal dual variables  $(\mu_1, \mu_2)$ , and the vector  $\bar{\lambda}$  of  $(\lambda_{kj})_{k=1, j=1}^K$ , where optimality here refers to the dual CBF optimization problem (22). Due to the dual feasibility constraints, the sequence of  $(\mu_1, \mu_2)$  is contained in a compact set and so the probability distribution function of  $\mu_1, F_1^{(N)}$ , forms a tight sequence [56]. Let  $F_1$  denote a limit point, so that  $F_1^{(N)} \Rightarrow F_1$  along a convergent subsequence.

For the purpose of obtaining a contradiction, let  $\bar{\mu}_1$  be such that  $F_1(\bar{\mu}_1 - \delta, \bar{\mu}_1 + \delta) > 0$  for all  $\delta > 0$ , and let  $\bar{\mu}_2 = 2 - \bar{\mu}_1$ . We will assume  $0 < \bar{\mu}_1 < 2$ .<sup>14</sup> Roughly speaking, there is non-negligible probability that  $\mu_1$  will be close to  $\bar{\mu}_1$  when  $N$  is large, along the subsequence. Indeed, for  $\delta > 0$ , let us define  $B_1(\delta)$  be the event that  $\mu_1 \in (\bar{\mu}_1 - \delta, \bar{\mu}_1 + \delta)$ , then by the second Borel-Cantelli lemma, event  $B_1(\delta)$  will occur infinitely often. Due to the feasibility constraint that  $\mu_1 + \mu_2 = 2$ , we can equivalently write

$$B_1(\delta) = \begin{cases} \bar{\mu}_1 - \delta \leq \mu_1 \leq \bar{\mu}_1 + \delta, \bar{\mu}_2 - \delta \leq \mu_2 \leq \bar{\mu}_2 + \delta & \delta > 0 \\ \bar{\mu}_1 + \delta \leq \mu_1 \leq \bar{\mu}_1 - \delta, \bar{\mu}_2 + \delta \leq \mu_2 \leq \bar{\mu}_2 - \delta & \delta < 0. \end{cases} \quad (169)$$

To compare the performance of the optimal scheme with a deterministic power scheme, we modify the SINR target in (160) from  $\gamma$  to  $\gamma + \delta$  for some small constant  $\delta$  (positive or negative), and we replace  $(\bar{\mu}_1, \bar{\mu}_2)$  by  $(\bar{\mu}_1 + \delta, \bar{\mu}_2 + \delta)$ . The latter change violates the constraint that the sum of the  $\mu$ s should be 2, but it nonetheless provides a valid pair of noise values for a virtual UL. Provided that  $|\delta|$  is sufficiently small, the condition  $0 < \bar{\mu}_i + \delta < 2$   $i = 1, 2$ , will be met.

Denote the corresponding solution to (161) (with  $(\mu_1, \mu_2)$  replaced by  $(\bar{\mu}_1 + \delta, \bar{\mu}_2 + \delta)$ ) by  $(\bar{\lambda}_1(\delta), \bar{\lambda}_2(\delta))$ , and let  $\bar{\lambda}(\delta)$  denote the vector of UL powers, where cell  $j$  users use power level  $\bar{\lambda}_j(\delta)$ . Note that  $(\bar{\lambda}_1(\delta), \bar{\lambda}_2(\delta))$  (and hence  $\bar{\lambda}(\delta)$ ) depends on  $(\bar{\mu}_1, \bar{\mu}_2)$ , and  $\delta$ . The analysis in Appendix B-B shows that this power allocation will asymptotically achieve a SINR of  $\gamma + \delta$  for all users, under external noise levels of  $\bar{\mu}_1 + \delta$ , and  $\bar{\mu}_2 + \delta$ , respectively, and for finite  $N$ , let it achieve SINR  $\gamma_{kj}(\delta)$  for user  $k$  in cell  $j$ .

As in the SCP case, Theorem 3 in [53] can be applied to show that for this (suboptimal) system, the value of (164) is

$$\gamma_{kj}(\delta) = \gamma + \delta + \mathcal{O} \left( \sqrt{\frac{1}{N}} \right), \quad \forall k = 1, 2, \dots, K, j = 1, 2 \quad (170)$$

where the last term on the right hand side is  $1/\sqrt{N}$  times a r.v. that tends weakly to a zero-mean Gaussian r.v. This holds since  $\frac{K}{N} \rightarrow \beta < \infty$ , for the considered channel model and in the notation of the theorem,  $\Gamma_{\bar{K}}$  and  $\Theta_{\bar{K}}$  will converge a.s. to bounded limits, though these will be different from the ones in the SCP case. As in the proof of Theorem 1 in Appendix D, a union bound may be applied to show that

$$\gamma_{kj}(-\delta) \leq \gamma \leq \gamma_{kj}(\delta) \quad \forall k = 1, 2, \dots, K, j = 1, 2 \quad (171)$$

will hold whenever  $B_1(\delta)$  occurs, in the limit as  $N \uparrow \infty$ .

Now compare the dual optimal power levels with this suboptimal, deterministic power allocation, on the event that  $B_1(\delta)$  occurs. Consider the case that  $\delta > 0$ : Noting (169), we see that, on this event, the dual UL powers required to achieve  $\gamma_{k,j}(\delta)$  will only decrease if we replace the enhanced noise levels,  $(\bar{\mu}_1 + \delta, \bar{\mu}_2 + \delta)$ , by  $(\mu_1, \mu_2)$ . But by (170) and the monotonicity result in Lemma 2 in Appendix C, it follows that these decreased UL powers must upper bound  $\bar{\lambda}$  since the latter is the optimal vector of powers in each cell to achieve

<sup>14</sup>The cases  $\bar{\mu}_1 = 0$  or  $\bar{\mu}_1 = 2$  can be considered separately, in a similar manner, which we mention at the conclusion of the proof.

the all- $\gamma$  SINR vector under noise levels  $(\mu_1, \mu_2)$ . Similarly, consider the case  $\delta < 0$ : On the event  $B_1(\delta)$ , the dual UL powers required to achieve  $\gamma_{k,j}(\delta)$ , will only increase if we replace the lower noise levels,  $(\bar{\mu}_1 + \delta, \bar{\mu}_2 + \delta)$ , by  $(\mu_1, \mu_2)$ , but by (170) and the monotonicity result in Lemma 2 in Appendix C, it follows that these increased UL powers lower bound  $\lambda$ . We conclude that for  $N$  sufficiently large, and on the event  $B_1(\delta)$ , we have that

$$\bar{\lambda}_1(-\delta) \leq \lambda_1 \leq \bar{\lambda}_1(\delta) \quad (172)$$

$$\bar{\lambda}_2(-\delta) \leq \lambda_2 \leq \bar{\lambda}_2(\delta) \quad (173)$$

and hence that the optimal dual objective value lies in the interval  $(\beta\sigma^2(\bar{\lambda}_1(-\delta) + \bar{\lambda}_2(-\delta)), \beta\sigma^2(\bar{\lambda}_1(\delta) + \bar{\lambda}_2(\delta)))$ . But  $\delta$  can be taken arbitrarily small, so the left and right endpoints of the interval can be made arbitrarily close to  $\beta\sigma^2(\bar{\lambda}_1 + \bar{\lambda}_2)$ , where  $(\bar{\lambda}_1, \bar{\lambda}_2)$  are the corresponding deterministic UL powers when we set  $\delta = 0$ : These are the unique solution to (161) for the given  $(\bar{\mu}_1, \bar{\mu}_2)$ . Thus, when  $N$  is sufficiently large, along the chosen subsequence, and on the event  $B_1(\delta)$  (which occurs infinitely often) we have the dual objective value getting as close as we like to the value  $\beta\sigma^2(\bar{\lambda}_1 + \bar{\lambda}_2)$ , which is a feasible value to the optimization problem (165)-(167). Thus if  $(\bar{\mu}_1, \bar{\mu}_2) \neq (1, 1)$  and  $0 < \bar{\mu}_1 < 2$ , then we get a contradiction of the optimality of the optimal dual power levels, whenever  $B_1(\delta)$  occurs, since we can beat the purported optimal value using the deterministic UL power levels obtained for  $(\mu_1, \mu_2) = (1, 1)$ , i.e. (45): this is the case since as discussed earlier the corresponding dual objective converges to the solution of (165)-(167). A similar argument can be used to show that  $(\bar{\mu}_1, \bar{\mu}_2) = (0, 2)$  or  $(\bar{\mu}_1, \bar{\mu}_2) = (2, 0)$  lead to similar contradictions. Since these are contradictions, it follows that  $(\bar{\mu}_1, \bar{\mu}_2) = (1, 1)$  and hence all the mass of the distribution  $F_1$  must be concentrated at  $(1, 1)$ .

We can conclude from the above analysis that for any positive  $\delta$ , the event  $B_1(\delta)$  will occur with probability tending to 1 as  $N$  tends to infinity, and therefore that (172)-(173) hold with probability also tending to 1. Taking  $\delta$  to zero, we obtain that the empirical distributions of both the  $\lambda_{1k}$ s and the  $\lambda_{2k}$ s tend to the same constant, namely  $\bar{\lambda}$ , as  $N$  tends to infinity, where  $\bar{\lambda}$  given in (45).

We conclude that the asymptotically optimal dual variables are  $\mu_1 = \mu_2 = 1$  and all  $\lambda_{kj}$ s converge to  $\bar{\lambda}$  as given in (45). Thus the optimal downlink beamforming vectors asymptotically point in the directions of the vectors given in (47). It remains to find the optimal power levels,  $p_{kj}$ , to use for each user.

#### Asymptotic analysis of the primal problem

As in Appendix D, we start by assuming that the  $p_{kj}$ 's in each cell are fixed at some common constant value  $\bar{p}_j$ , so that the total transmit power of cell  $j$  is  $\bar{P}_j = \beta\bar{p}_j$ . Once this regime is analyzed in the large system limit, we optimize the constants  $\bar{p}_1$  and  $\bar{p}_2$  (to meet the DL SINR constraints) and show that the optimal constants are asymptotically optimal with respect to the primal optimization problem (13).

Under this regime in which  $p_{kj} = \bar{p}_j$  for all  $k$  in cell  $j$ , the following lemmas hold.

**Lemma 7.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \mathbf{h}_{k,j}^H$ ,  $\bar{\lambda}$  satisfying (45),

$$\max_{j=1,2,k \leq K} \left| \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{kj}|^2}{N \|\hat{\mathbf{w}}_{kj}\|^2} - \left[ 1 - \beta \left( \frac{\gamma^2}{(1+\gamma)^2} + \frac{\epsilon^2 \gamma^2}{(1+\epsilon\gamma)^2} \right) \right] \right| \xrightarrow{a.s.} 0 \quad (174)$$

$$\max_{j=1,2,k,l \leq K} \left| \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{lj}|^2}{N \|\hat{\mathbf{w}}_{lj}\|^2} - \frac{\epsilon}{(1+\epsilon\gamma)^2} \right| \xrightarrow{a.s.} 0 \quad (175)$$

$$\max_{j=1,2,k,l \leq K, l \neq k} \left| \frac{|\mathbf{h}_{k,j} \hat{\mathbf{w}}_{lj}|^2}{N \|\hat{\mathbf{w}}_{lj}\|^2} - \frac{1}{(1+\gamma)^2} \right| \xrightarrow{a.s.} 0. \quad (176)$$

as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} \rightarrow \beta$ .

*Proof:* The proof is similar to that of Lemma 3, the main difference lying in the fact that there is only one case to consider when looking at the interference terms (cf. (135)), since under coordinated beamforming,  $\mathbf{h}_{k,j,\bar{j}}$  is no longer independent of  $\hat{\mathbf{w}}_{k',\bar{j}}$ . Thus, defining  $\mathbf{D}_{k,j,j} = \bar{\lambda} \mathbf{I}$  and  $\mathbf{D}_{k,j,\bar{j}} = \epsilon \bar{\lambda} \mathbf{I}$ , exactly as in the proof of Lemma 3, we can show that

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \|\hat{\mathbf{w}}_{kj}\|^2 - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_j^2 \right| \xrightarrow{a.s.} 0 \quad (177)$$

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j} \hat{\mathbf{w}}_{kj} - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_j \right| \xrightarrow{a.s.} 0. \quad (178)$$

When studying the interference terms, the numerator in the RHS of (135), becomes, for the CBf case,

$$\begin{aligned} & \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,i} \mathbf{A}_{l,i} \mathbf{h}_{l,i}^H \mathbf{h}_{l,i} \mathbf{A}_{l,i} \mathbf{h}_{k,j,i}^H \\ &= \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{l,i,i}^H \mathbf{h}_{l,i,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{k,j,i}^H \\ & \quad \left( 1 + \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{k,j,i}^H \right)^2. \end{aligned} \quad (179)$$

Considering the numerator in (179), we can show that

$$\begin{aligned} & \max_{j,i=1,2,k,l \leq K, (k,j) \neq (l,i)} \left| \frac{\bar{\lambda}^2}{N} \mathbf{h}_{k,j,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{l,i,i}^H \mathbf{h}_{l,i,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{k,j,i}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,i} \mathbf{A}_i \mathbf{D}_{l,i,i} \mathbf{A}_i \right| \xrightarrow{a.s.} 0. \end{aligned} \quad (180)$$

We can also show that

$$\begin{aligned} & \max_{j,i=1,2,k,l \leq K, (k,j) \neq (l,i)} \left| \frac{\bar{\lambda}}{N} \mathbf{h}_{k,j,i} \mathbf{A}_{l,i,k,j,i} \mathbf{h}_{k,j,i}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j,i} \mathbf{A}_i \right| \xrightarrow{a.s.} 0. \end{aligned} \quad (181)$$

The proof of the lemma is concluded by using the limits of the trace terms as derived in Appendix B-B, with  $\lambda_j = \bar{\lambda}$ , noting that for this specific case, all the  $\mathbf{D}_{k,j,j}$  and  $\mathbf{D}_{k,j,\bar{j}}$

matrices are equal, and are simply scaled identities, so that

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_j = \frac{\bar{\lambda}}{N} \text{tr} \mathbf{A}_j \xrightarrow{a.s.} \gamma \quad (182)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,\bar{j},j} \mathbf{A}_j = \frac{\epsilon \bar{\lambda}}{N} \text{tr} \mathbf{A}_j \xrightarrow{a.s.} \epsilon \gamma \quad (183)$$

$$\begin{aligned} \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_j^2 &= \frac{\bar{\lambda}}{N} \text{tr} \mathbf{A}_j^2 \\ &\xrightarrow{a.s.} \frac{1}{\bar{\lambda}} \frac{\gamma}{\frac{1}{\bar{\lambda}} + \frac{\beta}{(1+\gamma)^2} + \frac{\beta\epsilon}{(1+\epsilon\gamma)^2}} = \frac{1}{\bar{\lambda}} \frac{\gamma^2}{1 - \beta \left( \frac{\gamma^2}{(1+\gamma)^2} + \frac{\epsilon^2 \gamma^2}{(1+\epsilon\gamma)^2} \right)} \end{aligned} \quad (184)$$

$$\begin{aligned} \frac{1}{N} \text{tr} \mathbf{D}_{k,j,j} \mathbf{A}_j \mathbf{D}_{l,j,j} \mathbf{A}_j &= \frac{\bar{\lambda}^2}{N} \text{tr} \mathbf{A}_j^2 \\ &\xrightarrow{a.s.} \frac{\gamma^2}{1 - \beta \left( \frac{\gamma^2}{(1+\gamma)^2} + \frac{\epsilon^2 \gamma^2}{(1+\epsilon\gamma)^2} \right)} \end{aligned} \quad (185)$$

$$\begin{aligned} \frac{1}{N} \text{tr} \mathbf{D}_{k,\bar{j},j} \mathbf{A}_j \mathbf{D}_{l,j,j} \mathbf{A}_j &= \frac{\epsilon \bar{\lambda}^2}{N} \text{tr} \mathbf{A}_j^2 \\ &\xrightarrow{a.s.} \frac{\epsilon \gamma^2}{1 - \beta \left( \frac{\gamma^2}{(1+\gamma)^2} + \frac{\epsilon^2 \gamma^2}{(1+\epsilon\gamma)^2} \right)}. \end{aligned} \quad (186)$$

**Lemma 8.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \mathbf{h}_{k,j,j}^H$ , such that  $\bar{\lambda}$  satisfies (45), and with  $p_{kj} = \bar{p}_j$  for  $k = 1, \dots, K$ ,  $j = 1, 2$ , it follows that, with probability 1,

$$\sum_{j',k', (k',j') \neq (k,j)} \frac{p_{k'j'}}{N} \frac{|\mathbf{h}_{k,j,j'} \hat{\mathbf{w}}_{k'j'}|^2}{\|\hat{\mathbf{w}}_{k'j'}\|^2} \xrightarrow{a.s.} \frac{\bar{P}_j}{(1+\gamma)^2} + \frac{\epsilon \bar{P}_j}{(1+\epsilon\gamma)^2}, \quad (187)$$

$\bar{P}_j$  denotes the total transmit power of BS  $j$ , i.e.  $P_j = \beta \bar{P}_j$ ,  $j = 1, 2$ .

*Proof:* Since DL  $p_{kj}$ s are all fixed to the same value,  $\bar{p}_j$ , we have

$$\begin{aligned} &\sum_{j',k', (k',j') \neq (k,j)} \frac{p_{k'j'}}{N} \frac{|\mathbf{h}_{k,j,j'} \hat{\mathbf{w}}_{k'j'}|^2}{\|\hat{\mathbf{w}}_{k'j'}\|^2} \\ &= \frac{\bar{p}_j}{N} \sum_{k' \neq k} \frac{|\mathbf{h}_{k,j,j'} \hat{\mathbf{w}}_{k'j'}|^2}{\|\hat{\mathbf{w}}_{k'j'}\|^2} + \frac{\bar{p}_j}{N} \sum_{k'} \frac{|\mathbf{h}_{k,j,\bar{j}} \hat{\mathbf{w}}_{k'\bar{j}}|^2}{\|\hat{\mathbf{w}}_{k'\bar{j}}\|^2} \end{aligned} \quad (188)$$

Applying (176) and (175) of Lemma 7, (188) becomes for  $k = 1, \dots, K$ ,  $j = 1, 2$ ,

$$\frac{K-1}{N} \frac{\bar{p}_j}{(1+\gamma)^2} + \epsilon \beta \frac{\bar{p}_j}{(1+\epsilon\gamma)^2} + o(1). \quad (189)$$

Noting that as  $K, N \rightarrow \infty$ ,  $\frac{K-1}{N} \rightarrow \beta$  completes the proof. ■

Referring to (26), and using results from Lemmas 7 and 8, to find the *minimal* pair of constants  $(\bar{p}_1, \bar{p}_2)$  for the two cells,

we therefore solve the following set of equations,

$$\begin{aligned} \bar{p}_1 &\frac{1 - \frac{\beta\gamma^2}{(1+\gamma)^2} + \frac{\beta\epsilon^2\gamma^2}{(1+\epsilon\gamma)^2}}{\gamma} \\ &= \sigma^2 + \frac{\epsilon\beta}{(1+\epsilon\gamma)^2} \bar{p}_2 + \bar{p}_1 \frac{\beta}{(1+\gamma)^2}, \end{aligned} \quad (190)$$

$$\begin{aligned} \bar{p}_2 &\frac{1 - \frac{\beta\gamma^2}{(1+\gamma)^2} + \frac{\beta\epsilon^2\gamma^2}{(1+\epsilon\gamma)^2}}{\gamma} \\ &= \sigma^2 + \frac{\epsilon\beta}{(1+\epsilon\gamma)^2} \bar{p}_1 + \bar{p}_2 \frac{\beta}{(1+\gamma)^2}, \end{aligned} \quad (191)$$

to obtain

$$\bar{p}_1 = \bar{p}_2 = \bar{p} = \frac{\sigma^2 \gamma}{1 - \frac{\beta\gamma}{(1+\gamma)} - \epsilon \frac{\beta\gamma}{1+\epsilon\gamma}}. \quad (192)$$

This implies that

$$\bar{P}_1 = \bar{P}_2 = \bar{P} := \beta \bar{p}. \quad (193)$$

This choice of downlink transmit powers is guaranteed to meet the SINR constraints as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} = \beta$ .

We confirm the asymptotic optimality of this deterministic power allocation, together with the beamforming directions found from analysis of the dual problem, by verifying that the duality gap is tending to zero. Indeed, the primal objective value converges to  $2\bar{P}$  ( $\phi = 1$ , since the power consumption in both cells is the same), whereas the dual objective value converges to  $2\beta\bar{\lambda}\sigma^2$ . Recalling the definition of  $\bar{\lambda}$ , one can easily verify that these two quantities are the same, thereby completing the proof.

Equation (192) shows that the primal problem has a solution when effective bandwidth condition  $\beta \left( \frac{\gamma}{1+\gamma} + \frac{\epsilon\gamma}{1+\epsilon\gamma} \right) < 1$  is satisfied. Conversely, if this condition does not hold, there is asymptotically no feasible solution to the primal problem, in the limit as  $N \uparrow \infty$ . The latter observation follows from monotonicity: The optimal powers are increasing functions of  $\gamma$ , but as the denominator of (192) decreases to zero, the optimal total power from either BS,  $\bar{P}$ , must diverge to infinity, and for higher values of  $\gamma$  there can be no feasible solution.

## APPENDIX F PROOF OF THEOREM 3

Throughout this section, let

$$\mathbf{A} = \left[ \frac{\bar{\lambda}}{N} \sum_{j'=1}^2 \sum_{k'=1}^K \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{I} \right]^{-1} \quad (194)$$

$$\mathbf{A}_{k,j} = \left[ \frac{\bar{\lambda}}{N} \sum_{(k',j') \neq (k,j)} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{I} \right]^{-1} \quad (195)$$

$$\mathbf{A}_{k,j,l,i} = \left[ \frac{\bar{\lambda}}{N} \sum_{(k',j') \neq (k,j), (l,i)} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} + \mathbf{I} \right]^{-1} \quad (196)$$

where  $\bar{\lambda}$  will be specified later.

The following lemmas will be useful in proving the theorem.

**Lemma 9.** Let  $(\mu_1, \mu_2)$  satisfy  $0 \leq \mu_i \leq 2$ ,  $i = 1, 2$ , and define the function  $F(\lambda_1, \lambda_2)$  by

$$\begin{aligned} F(\lambda_1, \lambda_2) &= (F_1(\lambda_1, \lambda_2), F_2(\lambda_1, \lambda_2)) \\ F_j(\lambda_1, \lambda_2) &= \gamma \left[ \left( \mu_j + \frac{\beta}{1+\gamma} (\lambda_j + \lambda_{\bar{j}} \epsilon) \right)^{-1} \right. \\ &\quad \left. + \epsilon \left( \mu_{\bar{j}} + \frac{\beta}{1+\gamma} (\lambda_{\bar{j}} + \lambda_j \epsilon) \right)^{-1} \right]^{-1} \quad j = 1, 2. \end{aligned} \quad (197)$$

Then

- i)  $F$  is an interference function [27].
- ii) If  $\beta \frac{\gamma}{1+\gamma} < 1$  then there exists a unique solution  $(\bar{\lambda}_1, \bar{\lambda}_2)$  to the fixed point equation

$$(\lambda_1, \lambda_2) = F(\lambda_1, \lambda_2). \quad (198)$$

*Proof:* Similar to the proof of Lemma 6 with  $\tilde{\lambda}$  in (162) replaced by

$$\tilde{\lambda} = \frac{2\gamma}{1+\epsilon} \left( 1 - \beta \frac{\gamma}{1+\gamma} \right)^{-1}. \quad (199)$$

**Lemma 10.** Given positive  $(\mu_1, \mu_2)$ , and  $(\lambda_1, \lambda_2)$ , define function  $G(\gamma_1, \gamma_2)$  by

$$\begin{aligned} G(\gamma_1, \gamma_2) &= (G_1(\gamma_1, \gamma_2), G_2(\gamma_1, \gamma_2)) \\ G_j(\gamma_1, \gamma_2) &= \lambda_j \left( \left( \mu_j + \frac{\beta \lambda_j}{1+\gamma_j} + \frac{\epsilon \beta \lambda_{\bar{j}}}{1+\gamma_{\bar{j}}} \right)^{-1} \right. \\ &\quad \left. + \epsilon \left( \mu_{\bar{j}} + \frac{\beta \lambda_{\bar{j}}}{1+\gamma_{\bar{j}}} + \frac{\epsilon \beta \lambda_j}{1+\gamma_j} \right)^{-1} \right)^{-1} \quad j = 1, 2. \end{aligned} \quad (200)$$

Then

- i)  $G$  is an interference function [27].
- ii) If the inequalities

$$\gamma_1 \geq G_1(\gamma_1, \gamma_2), \quad \gamma_2 \geq G_2(\gamma_1, \gamma_2) \quad (201)$$

have a solution, then  $G$  has a unique fixed point.

*Proof:* (i) is easily verified, and (ii) follows from [27], Theorem 1. ■

**Corollary 5.** Let  $(\mu_1, \mu_2)$  satisfy  $0 \leq \mu_i \leq 2$ ,  $i = 1, 2$ . Assume that  $\beta \frac{\gamma}{1+\gamma} < 1$ , and let  $(\bar{\lambda}_1, \bar{\lambda}_2)$  be the unique fixed point in (197), as identified in Lemma 9. Then the function  $G$  in Lemma 10 has a unique fixed point, namely  $(\gamma, \gamma)$ .

*Asymptotic analysis of the dual problem*

Assume that  $\beta \frac{\gamma}{1+\gamma} < 1$ , let  $(\mu_1, \mu_2)$  be feasible for the dual (29) i.e.  $\mu_1, \mu_2 \geq 0$  and  $\mu_1 + \mu_2 = 2$ , and let  $(\bar{\lambda}_1, \bar{\lambda}_2)$  be the unique solution to (198). Consider the suboptimal dual power vector that assigns power level  $\bar{\lambda}_j/N$  to all the users in cell  $j$ ,  $j = 1, 2$ , respectively. The dual UL SINR for any user  $k$  in cell  $j$  is given by Equation (34). The derivations leading up to (102) in Appendix B-C show that with the given

suboptimal dual power vector assignment (34) will converge a.s. to  $(\gamma_1, \gamma_2)$  that is a fixed point of  $G$  as defined in Lemma 10. However, by Corollary 5,  $G$  has a unique fixed point, namely  $(\gamma, \gamma)$ . It follows that the SINRs of all users in the system tend to the common value  $\gamma$ .

We can follow the same reasoning as that in Appendix E, replacing the optimization problem (165) - (167) by:

$$\max_{\mu_1, \mu_2 \geq 0} \beta \sigma^2 (\bar{\lambda}_1 + \bar{\lambda}_2) \quad (202)$$

$$\text{s.t. } \mu_1 + \mu_2 = 2 \quad (203)$$

$$(\bar{\lambda}_1, \bar{\lambda}_2) \text{ is the unique fixed point of (198).} \quad (204)$$

The solution<sup>15</sup> to (202)-(204) can be shown to be  $(\mu_1, \mu_2) = (1, 1)$ , with corresponding  $(\bar{\lambda}_1, \bar{\lambda}_2) = (\bar{\lambda}, \bar{\lambda})$ , where  $\bar{\lambda}$  is given in (50); The details of the proof are skipped due to space constraints.

We thus conclude that the asymptotically optimal dual variables are  $\mu_1 = \mu_2 = 1$  and that all components of the optimal  $\lambda$  corresponding to cell  $j$  users must converge to  $\bar{\lambda}$ , as given in (50). Thus the optimal downlink beamforming vectors asymptotically point in the directions of the vectors given in (53). Fixing the beamforming to lie in these directions, we now find the optimal power levels,  $p_{kj}$ , to use for each user.

*Asymptotic analysis of the primal problem*

With the beamforming directions fixed, only the DL power levels  $p_{kj}$ 's still need to be determined. As in the proofs of Theorems 1 and 2, we first fix the DL powers to constants, and study the resulting limiting regime. Thus, we assume  $p_{kj}$ 's to be fixed at a constant  $\bar{p}_j$ ,  $j = 1, 2$ . Unlike the SCP and Cbf cases, this does not imply that the total transmit power of BS  $j$  is equal to  $\beta \bar{p}_j$ , but rather that the total transmit power of BS  $j$  is equal to

$$\bar{P}_j = \sum_{j'=1}^2 \frac{\bar{p}_{j'}}{N} \sum_{k=1}^K \frac{\|\mathbf{E}_j \hat{\mathbf{w}}_{kj'}\|^2}{\|\hat{\mathbf{w}}_{kj'}\|^2}. \quad (205)$$

Now turning to the DL power levels, we focus on the transmit strategy that allocates fixed power levels  $\bar{p}_j$  to all users in cell  $j$ . The following lemmas can be used to show that the left-hand side of (38) converges in probability to  $\sigma^2$  with all users being allocated equal power as given by (51). This power allocation also results in zero duality gap.

**Lemma 11.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \tilde{\mathbf{h}}_{k,j}^H$ , with  $\bar{\lambda}$  as given by (50):

$$\max_{j=1,2,k \leq K} \left| \frac{|\tilde{\mathbf{h}}_{k,j} \hat{\mathbf{w}}_{kj}|^2}{N \|\hat{\mathbf{w}}_{kj}\|^2} - (1+\epsilon) \left[ 1 - \frac{\beta \gamma^2}{(1+\gamma)^2} \right] \right| \rightarrow 0 \text{ a.s.} \quad (206)$$

$$\begin{aligned} &\max_{j=1,2,k,k' \leq K, k' \neq k} \left| \frac{|\tilde{\mathbf{h}}_{k',j} \hat{\mathbf{w}}_{kj}|^2}{\|\hat{\mathbf{w}}_{kj}\|^2} \right. \\ &\quad \left. - \frac{1}{1+\gamma^2} \left( 1 + \epsilon - \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right) \right| \rightarrow 0, \text{ i.p.} \end{aligned} \quad (207)$$

<sup>15</sup>Note that the function  $F$  in (197) actually depends on the choice of  $(\mu_1, \mu_2)$ .

$$\max_{j=1,2,k,k' \leq K, k' \neq k} \left| \frac{\tilde{\mathbf{h}}_{k',j} \tilde{\mathbf{w}}_{kj}}{\|\tilde{\mathbf{w}}_{kj}\|^2} \right|^2 - \frac{1}{1+\gamma^2} \frac{2\epsilon}{1+\epsilon} \frac{1}{1-\beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \rightarrow 0, \text{ i.p.} \quad (208)$$

as  $K, N \rightarrow \infty, \frac{K}{N} \rightarrow \beta$ .

*Proof:* The proof is similar to that of Lemmas 3 and 7. Thus, defining  $\mathbf{D}_{k,j}$  and  $\mathbf{D}_{k,j}$ , as in (103) and (104),  $\mu_1 = \mu_2 = 1$  and  $\lambda_1 = \lambda_2 = \bar{\lambda}$ , we can show, by applying Lemma 5.1 in [55] and Lemma 1 that

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \|\tilde{\mathbf{w}}_{kj}\|^2 - \frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A}^2 \right| \xrightarrow{a.s.} 0 \quad (209)$$

$$\max_{j=1,2,k \leq K} \left| \frac{\bar{\lambda}}{N} \tilde{\mathbf{h}}_{k,j} \tilde{\mathbf{w}}_{kj} - \frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \right| \xrightarrow{a.s.} 0. \quad (210)$$

Now consider the interference terms,  $\frac{|\tilde{\mathbf{h}}_{k',j} \tilde{\mathbf{w}}_{k'j'}|^2}{\|\tilde{\mathbf{w}}_{k'j'}\|^2}$ , with  $(k', j') \neq (k, j)$ . We have that

$$\begin{aligned} & \frac{|\tilde{\mathbf{h}}_{k',j} \tilde{\mathbf{w}}_{k'j'}|^2}{\|\tilde{\mathbf{w}}_{k'j'}\|^2} \\ &= \frac{1}{\bar{\lambda}} \frac{\tilde{\mathbf{h}}_{k,j}^2 \mathbf{A}_{k',j',k,j} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} \mathbf{A}_{k',j',k,j} \tilde{\mathbf{h}}_{k,j}^H}{\left(1 + \frac{\bar{\lambda}}{N} \tilde{\mathbf{h}}_{k,j} \mathbf{A}_{k',j',k,j} \tilde{\mathbf{h}}_{k,j}^H\right)^2} \frac{1}{\bar{\lambda} \|\tilde{\mathbf{w}}_{k'j'}\|^2} \end{aligned} \quad (211)$$

Similarly to (209) and (210), we can show that

$$\begin{aligned} & \max_{j,i=1,2,k,l \leq K, (k,j) \neq (l,i)} \left| \frac{\bar{\lambda}^2}{N} \tilde{\mathbf{h}}_{k,j} \mathbf{A}_{k',j',k,j} \tilde{\mathbf{h}}_{k',j'}^H \tilde{\mathbf{h}}_{k',j'} \mathbf{A}_{k',j',k,j} \tilde{\mathbf{h}}_{k,j}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \mathbf{D}_{l,i} \mathbf{A} \right| \xrightarrow{a.s.} 0. \quad (212) \end{aligned}$$

We can also show that

$$\begin{aligned} & \max_{j,i=1,2,k,l \leq K, (k,j) \neq (l,i)} \left| \frac{\bar{\lambda}}{N} \tilde{\mathbf{h}}_{k,j} \mathbf{A}_{l,i,k,j} \tilde{\mathbf{h}}_{k,j}^H \right. \\ & \quad \left. - \frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \right| \xrightarrow{a.s.} 0. \quad (213) \end{aligned}$$

The proof of the lemma is concluded by using the limits of the trace terms as derived in Appendix B-C, with  $\lambda_j = \bar{\lambda}$  given by (50),

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \xrightarrow{a.s.} \gamma \quad (214)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A}^2 \xrightarrow{a.s.} \frac{1}{(1+\epsilon)\bar{\lambda}} \frac{\gamma^2}{1 - \frac{\beta\gamma^2}{(1+\gamma)^2}} \quad (215)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \mathbf{D}_{l,i} \mathbf{A} \xrightarrow{i.p.} \gamma^2 \frac{\frac{1+\epsilon^2}{(1+\epsilon)^2} - \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\beta\gamma^2}{(1+\gamma)^2}}{\left[1 - \frac{\beta\gamma^2}{(1+\gamma)^2}\right] \left[1 - \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\beta\gamma^2}{(1+\gamma)^2}\right]} \quad (216)$$

$$\frac{1}{N} \text{tr} \mathbf{D}_{k,j} \mathbf{A} \mathbf{D}_{l,i} \mathbf{A} \xrightarrow{i.p.} \gamma^2 \frac{\frac{2\epsilon}{(1+\epsilon)^2}}{\left[1 - \frac{\beta\gamma^2}{(1+\gamma)^2}\right] \left[1 - \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\beta\gamma^2}{(1+\gamma)^2}\right]}. \quad (217)$$

**Lemma 12.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \tilde{\mathbf{h}}_{k,j}^H$ , with  $\bar{\lambda}$  as given by (50), as  $K, N \rightarrow \infty, \frac{K}{N} \rightarrow \beta$ , the following holds with probability 1, for  $j = 1, 2$ ,

$$\bar{P}_j \rightarrow \sum_{j'=1}^2 \frac{\beta \bar{p}_{j'}}{2}, \quad (218)$$

where  $\bar{P}_j$  is the total transmit power under the considered DL transmit power regime (see (205)).

*Proof:* Following similar steps to those in the proof of Lemma 11, we can show that the  $\frac{\bar{\lambda}}{N} \|\mathbf{E}_j \tilde{\mathbf{w}}_{kj}\|^2$  and  $\frac{\bar{\lambda}}{N} \|\tilde{\mathbf{w}}_{kj}\|^2$  converge a.s., and uniformly in their indices, to constants such that the following holds

$$\max_{j,i=1,2,k \leq K} \left| \frac{\|\mathbf{E}_j \tilde{\mathbf{w}}_{kj}\|^2}{\|\tilde{\mathbf{w}}_{kj}\|^2} - \frac{1}{2} \right| \xrightarrow{a.s.} 0. \quad (219)$$

As a result,

$$\bar{P}_j = \sum_{j'=1}^2 \frac{\beta \bar{p}_{j'}}{2} + o(1). \quad (220)$$

This completes the proof. ■

Thus, asymptotically, under this scheme, both BSs transmit with equal power.

**Lemma 13.** With  $\hat{\mathbf{w}}_{kj} = \mathbf{A}_{k,j} \tilde{\mathbf{h}}_{k,j}^H$ , such that  $\bar{\lambda}$  satisfies (50), and with  $p_{kj} = \bar{p}_j$  for  $k = 1, \dots, K, j = 1, 2$ , it follows that,

$$\begin{aligned} & \sum_{j',k', (k',j') \neq (k,j)} \frac{p_{k'j'}}{N} \frac{|\tilde{\mathbf{h}}_{k,j} \tilde{\mathbf{w}}_{k'j'}|^2}{\|\tilde{\mathbf{w}}_{k'j'}\|^2} \\ & \xrightarrow{i.p.} \frac{\beta}{1+\gamma^2} \left[ \bar{p}_j \left( 1 + \epsilon - \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right) \right. \\ & \quad \left. + \bar{p}_{\bar{j}} \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right]. \quad (221) \end{aligned}$$

*Proof:* Since DL  $p_{kjs}$  are all fixed to the same value,  $\bar{p}_j$ , we have

$$\begin{aligned} & \sum_{j',k', (k',j') \neq (k,j)} \frac{p_{k'j'}}{N} \frac{|\tilde{\mathbf{h}}_{k,j} \tilde{\mathbf{w}}_{k'j'}|^2}{\|\tilde{\mathbf{w}}_{k'j'}\|^2} \\ &= \frac{\bar{p}_j}{N} \sum_{k' \neq k} \frac{|\tilde{\mathbf{h}}_{k,j} \tilde{\mathbf{w}}_{k'j}|^2}{\|\tilde{\mathbf{w}}_{k'j}\|^2} + \frac{\bar{p}_{\bar{j}}}{N} \sum_{k'} \frac{|\tilde{\mathbf{h}}_{k,j} \tilde{\mathbf{w}}_{k'j}|^2}{\|\tilde{\mathbf{w}}_{k'j}\|^2} \quad (222) \end{aligned}$$

Applying (207) and (208) of Lemma 7, (222) for  $k = 1, \dots, K, j = 1, 2$ , will converge in probability to

$$\begin{aligned} & \frac{1}{1+\gamma^2} \left[ \bar{p}_j \frac{K-1}{N} \left( 1 + \epsilon - \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right) \right. \\ & \quad \left. + \bar{p}_{\bar{j}} \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right]. \quad (223) \end{aligned}$$

Noting that as  $K, N \rightarrow \infty, \frac{K-1}{N} \rightarrow \beta$  completes the proof. ■

Referring to (38), and applying Lemmas 11 and 13, to find the *minimal* pair of constants  $(\bar{p}_1, \bar{p}_2)$  for users in the two cells, we therefore solve the following set of equations,

$$\begin{aligned} & \bar{p}_1 \frac{(1+\epsilon) \left[ 1 - \frac{\beta\gamma^2}{(1+\gamma)^2} \right]}{\gamma} \\ &= \sigma^2 + \frac{\beta}{1+\gamma^2} \left[ \bar{p}_1 \left( 1 + \epsilon - \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right) \right. \\ & \quad \left. + \bar{p}_2 \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right], \end{aligned} \quad (224)$$

$$\begin{aligned} & \bar{p}_2 \frac{(1+\epsilon) \left[ 1 - \frac{\beta\gamma^2}{(1+\gamma)^2} \right]}{\gamma} \\ &= \sigma^2 + \frac{\beta}{1+\gamma^2} \left[ \bar{p}_2 \left( 1 + \epsilon - \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right) \right. \\ & \quad \left. + \bar{p}_1 \frac{\frac{2\epsilon}{1+\epsilon}}{1 - \beta \frac{(1-\epsilon)^2}{(1+\epsilon)^2} \frac{\gamma^2}{(1+\gamma)^2}} \right], \end{aligned} \quad (225)$$

to obtain

$$\bar{p}_1 = \bar{p}_2 = \bar{p} = \frac{\sigma^2 \gamma}{(1+\epsilon) \left[ 1 - \frac{\beta\gamma}{1+\gamma} \right]}. \quad (226)$$

This implies that  $\bar{P}_1$  and  $\bar{P}_2$  (recall Lemma 12) both converge a.s. to  $\bar{P}$ , where

$$\bar{P} := \beta \bar{p}. \quad (227)$$

This choice of DL transmit powers meets the SINR constraints as  $K, N \rightarrow \infty$ ,  $\frac{K}{N} = \beta$  with a probability tending to 1.

We confirm the asymptotic optimality of this deterministic power allocation, together with the beamforming directions found from analysis of the dual problem, by verifying that the duality gap is tending to zero. Indeed, the primal objective value converges to  $2\bar{P}$  ( $\phi = 1$ , since the power consumption in both cells is the same), whereas the dual objective value converges to  $2\beta\bar{\lambda}\sigma^2$ . Recalling the definition of  $\bar{\lambda}$ , one can easily verify that these two quantities are the same, thereby completing the proof.

Equation (226) shows that the primal problem have a solution when effective bandwidth condition  $\beta \left( \frac{\gamma}{1+\gamma} \right) < 1$  is satisfied. Conversely, if this condition does not hold, there is asymptotically no feasible solution to the primal problem, in the limit as  $N \uparrow \infty$ .

#### APPENDIX G PROOF OF THEOREM 1

Using (43),  $r$  simplifies to:

$$r(\gamma^*) = \frac{1 + \gamma^*}{\gamma^* \left( \frac{\sigma^2}{P} + \epsilon \right) (1 + \gamma^*) + \gamma^*} \log(1 + \gamma^*), \quad (228)$$

which is positive at  $\gamma^* = 0$ , and zero at  $\gamma^* = \infty$ . Defining  $\eta = \frac{\sigma^2}{P} + \epsilon$  and taking the derivative with respect to  $\gamma^*$  we obtain  $\frac{dr}{d\gamma^*}$  equals

$$-\frac{1 + \eta(1 + \gamma^*)^2}{(\gamma^* \eta(1 + \gamma^*) + \gamma^*)^2} \log(1 + \gamma^*) + \frac{1}{\gamma^* \eta(1 + \gamma^*) + \gamma^*}. \quad (229)$$

This has the same sign as

$$h(\gamma^*) = -\log(1 + \gamma^*) + \frac{\gamma^* \eta(1 + \gamma^*) + \gamma^*}{1 + \eta(1 + \gamma^*)^2}, \quad (230)$$

which is 0 at  $\gamma^* = 0$ , and  $-\infty$  at  $\gamma^* = \infty$ . Differentiating the expression in (230), we obtain

$$\frac{dh}{d\gamma^*} = -\frac{1}{1 + \gamma^*} + \frac{\eta + 1 + \eta(\eta + 1)(1 + \gamma^*)^2 - 2\eta\gamma^{*2}}{(1 + \eta(1 + \gamma^*)^2)^2}. \quad (231)$$

This has the same sign as

$$\begin{aligned} & -\left(1 + \eta(1 + \gamma^*)^2\right)^2 \\ & + (1 + \gamma^*) \left( \eta + 1 + (\eta + 1)\eta(1 + \gamma^*)^2 - 2\eta\gamma^{*2} \right) \\ & = -\gamma^{*4}\eta^2 - \gamma^{*3}\eta[1 + 3\eta] - \gamma^{*2}\eta[1 + 3\eta] - \gamma^*[\eta^2 - 1]. \end{aligned}$$

If  $\eta \geq 1$ , then  $\frac{dh}{d\gamma^*} < 0 \quad \forall \gamma^* > 0$ , which implies  $h(\gamma^*) < 0 \quad \forall \gamma^* > 0$  and hence  $r(\gamma^*)$  is a decreasing function. If  $\eta < 1$  then  $h(\gamma^*)$  is positive for  $\gamma^*$  small. This implies that  $r(\gamma^*)$  is increasing and then decreasing.

#### APPENDIX H PROOF OF THEOREM 2

The normalized achievable rate per cell is equal to

$$\begin{aligned} r(\gamma^*) &= \beta \log(1 + \gamma^*) = \frac{1}{\gamma^*} \frac{\log(1 + \gamma^*)}{\left[ \frac{\sigma^2}{P} + \frac{1}{1+\gamma^*} + \frac{\epsilon}{1+\epsilon\gamma^*} \right]} \\ &= \frac{(1 + \gamma^*)(1 + \epsilon\gamma^*) \log(1 + \gamma^*)}{[\gamma^* \left[ \frac{\sigma^2}{P} (1 + \gamma^*) + 1 \right] (1 + \epsilon\gamma^*) + \epsilon\gamma^* (1 + \gamma^*)]}, \end{aligned} \quad (232)$$

which is positive at  $\gamma^* = 0$  and zero at  $\gamma^* = \infty$ . The sign of  $\frac{dr}{d\gamma^*}$  is the same as the sign of the function

$$\begin{aligned} & h(\gamma^*) \\ &= \frac{(1 + \epsilon\gamma^*) \left[ \gamma^* \left[ \frac{\sigma^2}{P} (1 + \gamma^*) + 1 \right] (1 + \epsilon\gamma^*) + \epsilon\gamma^* (1 + \gamma^*) \right]}{\left[ (\epsilon\gamma^{*2} + 1) [1 + \epsilon] + \frac{\sigma^2}{P} (1 + \gamma^*)^2 (1 + \epsilon\gamma^*)^2 + 4\epsilon\gamma^* \right]} \\ & - \log(1 + \gamma^*) \end{aligned} \quad (233)$$

which is 0 at  $\gamma^* = 0$ , and  $-\infty$  at  $\gamma^* = \infty$ . The sign of  $\frac{dh}{d\gamma^*}$  can be shown to be *opposite* to the sign of a polynomial in  $\gamma^*$  of degree 7, i.e.  $\sum_{i=0}^7 c_i \gamma^{*i}$ , such that

$$\begin{aligned} c_7 &= a^2 \epsilon^4, \quad c_6 = \epsilon^3 (3a^2 \epsilon + 4a^2 + 2a\epsilon) \\ c_5 &= \epsilon^2 (3a^2 \epsilon^2 + 12a^2 \epsilon + 6a^2 + 4a\epsilon^2 + 6a\epsilon) \\ c_4 &= \epsilon (a^2 [\epsilon(\epsilon^2 + 12\epsilon + 18) + 4] + a\epsilon (3\epsilon^2 + 8\epsilon + 9) \\ & \quad - 2\epsilon(\epsilon + 1)) \end{aligned}$$



$$\begin{aligned}
c_3 &= a^2 [2\epsilon (2\epsilon^2 + 9\epsilon + 6) + 1] + a\epsilon (\epsilon^3 + 13\epsilon + 6) \\
&\quad + \epsilon^2 (1 - \epsilon^2 - 16\epsilon) \\
c_2 &= 3a^2 [2\epsilon (\epsilon + 2) + 1] + a (3\epsilon^2 + 10\epsilon + 1) + \epsilon \\
&\quad - (17\epsilon^3 + 8\epsilon^2 + 4a\epsilon^3) \\
c_1 &= a^2 (4\epsilon + 3) + 6a\epsilon + a - \epsilon (2a\epsilon^2 + 3a\epsilon + 9\epsilon^2 + 4\epsilon + 3) \\
c_0 &= a^2 + 2a\epsilon - 2a\epsilon^2 - 2\epsilon^3 - \epsilon^2 - 1,
\end{aligned} \tag{234}$$

where  $a = \frac{\sigma^2}{P}$ . One can factor the constant term (independent of  $\gamma^*$ ) into

$$(a + \epsilon + 1)(a + \epsilon - 2\epsilon^2 - 1). \tag{235}$$

We now consider two cases:

A.  $a + \epsilon - 2\epsilon^2 - 1 \geq 0$

Defining  $\epsilon_{0,1} = \frac{1-\sqrt{8a-7}}{4}$ ,  $\epsilon_{0,2} = \frac{1+\sqrt{8a-7}}{4}$ , one can show that (235) will be positive if ( $a \geq 1$ , and  $0 \leq \epsilon \leq \epsilon_{0,2}$ ), or ( $\frac{7}{8} \leq a \leq 1$ , and  $\epsilon_{0,1} \leq \epsilon \leq \epsilon_{0,2}$ ). One can further establish that if this is the case, all of the other coefficients of the polynomial (234) will be positive<sup>16</sup>, and consequently  $h(\gamma^*)$  will be a decreasing function, and hence  $h(\gamma^*) < 0 \quad \forall \gamma^* > 0$ . In this case,  $r(\gamma^*)$  is a decreasing function and the optimum will occur at  $\gamma^* = 0$ , which corresponds to  $\beta^* = \infty$ .

B.  $a + \epsilon - 2\epsilon^2 - 1 < 0$

In this case,  $h(\gamma^*)$  will be positive for  $\gamma^*$  small. Thus,  $r(\gamma^*)$  will be increasing initially, but it decreases eventually (to zero), so it must attain its global maximum at a finite value of  $\gamma^* > 0$ .

## REFERENCES

- [1] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, Sept. 2006.
- [2] H. Dahrouj and W. Yu, "Coordinated beamforming for the multi-cell multi-antenna wireless system," in *Proc. Conference on Information Sciences and Systems 2008, CISS 2008*, 2008.
- [3] H. Dahrouj and W. Yu, "Coordinated beamforming for the multicell, multi-antenna wireless system," *IEEE Trans. Wireless Commun.*, May 2010.
- [4] K. Karakayali et al., "On the maximum common rate achievable in a coordinated network," in *Proc. IEEE International Conference on Communications, ICC*, 2006.
- [5] S. Jing et al., "Multicell downlink capacity with coordinated processing," *EURASIP Journal on Wireless Communications and Networking*, 2008.
- [6] O. Somekh et al., "Cooperative multicell zero-forcing beamforming in cellular downlink channels," *IEEE Trans. Inf. Theory*, July 2009.
- [7] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE Vehicular Technology Conf.*, 2001.
- [8] R. Zakhour and S.V. Hanly, "Minmax coordinated beamforming via large system analysis," 2011, IEEE ISIT 2011.
- [9] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, 2004.
- [10] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, Mar. 1999.
- [11] Huaiyu Dai and H. V. Poor, "Asymptotic spectral efficiency of multicell MIMO systems with frequency-flat fading," *IEEE Trans. Signal Processing*, Nov. 2003.
- [12] A.M. Tulino, A. Lozano, and S. Verdu, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, July 2005.
- [13] D. Aktas et al., "Scaling results on the sum capacity of cellular networks with MIMO links," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, July 2006.
- [14] J. Dumont et al., "On the capacity achieving covariance matrix for Rician MIMO channels: An asymptotic approach," *IEEE Trans. Inf. Theory*, Mar. 2010.
- [15] V. K. Nguyen and J. S. Evans, "Multiuser transmit beamforming via regularized channel inversion: A large system analysis," in *Proc. IEEE Global Telecomm. Conf.*, Nov.-Dec. 2008.
- [16] R. Couillet et al., "The space frontier: Physical limits of multiple antenna information transfer," in *Proc. of VALUETOOLS 2008, Workshop Inter-Perf*, Oct. 2008.
- [17] R. Couillet, M. Debbah, and J. W. Silverstein, "A deterministic equivalent for the capacity analysis of correlated multi-user MIMO channels," *CoRR*, vol. abs/0906.3667, 2009.
- [18] S.H. Moon et al., "Weighted sum rate of multi-cell MIMO downlink channels in the large system limit," in *Proc. IEEE Int. Conf. on Communication*, June 2010.
- [19] H. Huh et al., "Network MIMO large-system analysis and the impact of CSIT estimation," Mar. 2010.
- [20] H. Huh et al., "Multi-cell MIMO downlink with fairness criteria: the large system limit," in *Proc. IEEE Int. Symp. on Inf. Theory*, June 2010.
- [21] S. Lakshminarayana, J. Hoydis, M. Debbah, and M. Assaad, "Asymptotic analysis of distributed multi-cell beamforming," in *Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications*, Sept. 2010.
- [22] R. Zakhour and S. V. Hanly, "Large system analysis of base station cooperation on the downlink," in *Proc. Allerton Conf. 2010*, IL, United States, Sept.-Oct. 2010.
- [23] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large system analysis of linear precoding in MISO broadcast channels with limited feedback," 2010, arXiv:0906.3682v3.
- [24] J. Hoydis, M. Kobayashi, and M. Debbah, "On the optimal number of cooperative base stations in network MIMO systems," 2010, arXiv:0912.4595v3.2642.
- [25] E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. Telecomm.*, Nov. 1999.
- [26] F. Rashid-Farrokhi, K. J. R. Liu, and L. Tassiulas, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, Oct. 1998.
- [27] R.D. Yates, "A framework for uplink power control in cellular radio networks," *IEEE J. Sel. Areas Commun.*, Sept. 1995.
- [28] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Vehicular Technology Conf.*, July 1999.
- [29] P. Viswanath and D.N.C. Tse, "Sum capacity of the multiple antenna Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, Aug. 2003.
- [30] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, Oct. 2003.
- [31] W. Yu and J. Cioffi, "Sum capacity of Gaussian vector broadcast channels," *IEEE Trans. Inf. Theory*, Sept. 2004.
- [32] Wei Yu, "Uplink-downlink duality via minimax duality," *IEEE Trans. Inf. Theory*, Feb. 2006.
- [33] D. N. C. Tse and P. Viswanath, "Downlink-uplink duality and effective bandwidths," in *Proc. IEEE Int. Symp. on Inf. Theory*, 2002.
- [34] H. Boche and M. Schubert, "A general duality theory for uplink and downlink beamforming," in *Proc. IEEE Vehicular Technology Conf.*, Sept. 2002.
- [35] B. R. Vojcic and W. M. Jang, "Transmitter precoding in synchronous multiuser communications," *IEEE Trans. Commun.*, 1998.
- [36] A.N. Baretto and G. Fettweis, "Capacity increase in the downlink of spread spectrum systems through joint signal processing," in *Proc. IEEE Int. Conf. on Communication*, Jun. 2001, vol. 4.
- [37] R. L.-U Choi and R. D. Murch, "Transmit-precoding techniques with simplified receivers for the downlink of MISO TDD-CDMA systems," *IEEE Trans. Vehicular Technology*, 2004.
- [38] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communication systems," *IEEE Trans. Signal Processing*, 2005.
- [39] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Processing*, vol. 54, no. 1, Jan. 2006.
- [40] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vehicular Technology*, Jan. 2004.

<sup>16</sup>This is done by using the fact that if  $a \geq \frac{7}{8}$  and  $0 \leq \epsilon \leq \epsilon_{0,2}$  then  $\epsilon \leq \sqrt{a}$  and  $\epsilon \leq a$ .

- [41] H. Boche and M. Schubert, "Effective bandwidth maximization for uplink/downlink multi-antenna systems," in *Proc. IEEE Int. Conf. on Communication*, May 2003, vol. 5.
- [42] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Processing*, June 2007.
- [43] Miljanic Z. Foschini G.J., "A simple distributed autonomous power control algorithm and its convergence," *IEEE Trans. Vehicular Technology*, no. 4, 1993.
- [44] C. B. Peel et al., "A vector-perturbation technique for near-capacity multiantenna multiuser communication-part I: channel inversion and regularization," *IEEE Trans. Commun.*, Jan. 2005.
- [45] A. Abramovich, "A controlled method for adaptive optimization of filters using the criterion of maximum snr," *Radiotekh. Elektron*, vol. 26, no. 3, 1981.
- [46] X. Mestre and M. A. Lagunas, "Finite sample size effect on minimum variance beamformers: optimum diagonal loading factor for large arrays," *IEEE Trans. Signal Processing*, vol. 54, no. 1, 2006.
- [47] M. K. Karakayali, G. J. Foschini, and R. A. Valenzuela, "Network coordination for spectrally efficient communications in cellular systems," *IEEE Wireless Communications*, vol. 13, Aug. 2006.
- [48] D. Gesbert et al., "Multicell MIMO cooperative networks: a new look at interference," *IEEE Journal on Selected Areas*, Dec. 2010.
- [49] A. Wyner, "Shannon-theoretic approach to a gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, 1994.
- [50] S. V. Hanly and D. N. C. Tse, "Resource pooling and effective bandwidths in CDMA networks with multiuser receivers and spatial diversity," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, May 2001.
- [51] S.V. Hanly, "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, Sept. 1995.
- [52] W. Hachem, P. Loubaton, and J. Najim, "A clt for information-theoretic statistics of gram random matrices with a given variance profile," *The Annals of Applied Probability*, vol. 18, no. 6, 2008.
- [53] A. Kammoun, M. Kharouf, W. Hachem, and J. Najim, "A central limit theorem for the SINR at the LMMSE estimator output for large-dimensional signals," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, Nov. 2009.
- [54] W. Hachem, P. Loubaton, and J. Najim, "Deterministic equivalents for certain functionals of large random matrices," *The Annals of Applied Probability*, vol. 17, no. 3, 2007.
- [55] Liang Ying-Chang, Pan Guangming, and Z. D. Bai, "Asymptotic performance of MMSE receivers for large systems using random matrix theory," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4173, 2007.
- [56] P. Billingsley, *Convergence of Probability Measures*, 2nd ed., John Wiley and Sons, 1999.